DOCUMENT RESUME

ED 176 220 CS 004 975

AUTHOR	Johnson, Dale D.; And Others
TITLE	Relationships between Word Identification Skills and Reading Comprehension of Elementary School Children.
INSTITUTION	Wisconsin Univ., Madison. Research and Development Center for Individualized Schooling.
SPONS AGENCY	National Inst. of Education (DHEW), Washington, D.C.
REPORT NO	WRDCIS-TR-490
PUB DATE	Dec 78
GRANT	OB-NIE-G-78-0217
NOTE	123p.
EDRS PRICE	MF01/PC05 Plus Postage.
DESCRIPTORS	*Context Clues; *Decoding (Reading); Elementary Education; Evaluation Methods; Phonics; *Reading Comprehension; Reading Diagnosis; *Reading Research; Reading Tests; Structural Analysis; *Word Recognition

ABSTRACT

Three studies were conducted as part of a series designed to identify the subskills of word identification that correlate most highly with reading comprehension and to examine various methods of assessing these subskills. In the first study, the prototype of a decoding (word identification) test was administered to 282 first, third, and fifth grade students, along with The Reading Subtest of the Metropolitan Achievement Tests (Form F). Study Two tested 792 second, fourth, and sixth grade students with a revised version of the decoding test and the Reading Subtest of the MAT. After another revision of the decoding test, the same procedures were used in Study Three to test 360 second, fourth, and sixth grade students. Results from the three studies indicated that all three broad areas of word identification--phonics, structural analysis, and contextual analysis--are significantly related to comprehension. The contextual analysis subtests in each study achieved higher correlation overall with the comprehension measure than did the structure and phonics subtests. Changes in the subtests between Studies Two and Three generally resulted in higher reliability ratings for those subtests that were extensively revised. The revision of subtests, however, made comparisons of results across the three studies inappropriate, and accounts for the considerable variation of correlations between skill subtests and global comprehension measures. (Author/RI)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED176220

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

Technical Report No. 490

RELATIONSHIPS BETWEEN WORD IDENTIFICATION SKILLS AND READING COMPREHENSION OF ELEMENTARY SCHOOL CHILDREN

by

Dale D. Johnson, Susan D. Pittelman, Judy Schwenker,
Linda K. Shriberg, and Carol Morgan-Janty

Report from the Project on
Studies in Language: Reading and Communication

Dale D. Johnson
Faculty Associate

CS204175

Wisconsin Research and Development Center
for Individualized Schooling
The University of Wisconsin
Madison, Wisconsin

December 1978

Published by the Wisconsin Research and Development Center for Individualized Schooling. The project presented or reported herein was performed pursuant to a grant from the National Institute of Education, Department of Health, Education, and Welfare. However, the opinions expressed herein do not necessarily reflect the position or policy of the National Institute of Education, and no official endorsement by the National Institute of Education should be inferred.

Center Grant No. OB-NIE-G-78-0217

MISSION STATEMENT

The mission of the Wisconsin Research and Development Center is to improve the quality of education by addressing the full range of issues and problems related to individualized schooling. Teaching, learning, and the problems of individualization are given concurrent attention in the Center's efforts to discover processes and develop strategies and materials for use in the schools. The Center pursues its mission by

- conducting and synthesizing research to clarify the processes of school-age children's learning and development
- conducting and synthesizing research to clarify effective approaches to teaching students basic skills and concepts
- developing and demonstrating improved instructional strategies, processes, and materials for students, teachers, and school administrators
- providing assistance to educators which helps transfer the outcomes of research and development to improved practice in local schools and teacher education institutions

The Wisconsin Research and Development Center is supported with funds from the National Institute of Education and the University of Wisconsin.

WISCONSIN RESEARCH AND DEVELOPMENT
CENTER FOR INDIVIDUALIZED SCHOOLING

Table of Contents

	Page
List of Tables	vi
List of Figures	ix
Abstract	xi
Introduction	1
Study 1	7
Method	7
Subjects	7
Stimuli	7
Procedure	9
Results	11
Summary	16
Study 2	19
Method	19
Subjects	19
Stimuli	21
Procedure	22
Results	25
Summary	40
Study 3	43
Method	43
Subjects	43
Stimuli	44
Procedure	47
Results	51
Summary	68
Conclusion	71
Reference Notes	77
References	79
Appendices	
A. A Description of the Decoding Test by Grade Level, Spring 1977; Sample Items from Each Subtest of the Decoding Test, Spring 1977	81
B. A Description of the Word Identification Test, Winter 1977-78; Sample Items from Each Subtest of the Word Identification Test, Winter 1977-78	87
C. A Description of the Word Identification Test, Spring 1978; Sample Items from Each Subtest of the Word Identification Test, Spring 1978	103
D. Teacher Questionnaire, Beginning Level; Teacher Questionnaire, Intermediate Level; Teacher Questionnaire, Advanced Level	115

List of Tables

Table		Page
1	Subjects by Grade Level and School District	8
2	Decoding Test: Subtests and Number of Items by Grade Level	10
3	Mean Scores (Raw and Percent Correct) by Grade Level	12
4	Pearson Correlations Between Overall Decoding and Comprehension, by Grade Level	13
5	Correlation Matrix for Decoding Subtests and Metropolitan Subtest (Comprehension)	14
6	Subjects by Grade Level and School District	20
7	Word Identification Test: Subtests and Number of Items by Grade Level and Form	23
8	Mean Scores (Raw and Percent Correct) by Grade Level	26
9	Hoyt Reliability Coefficients for Word Identification Subtests	28
10	Pearson Correlations of Word Identification Subtests with Metropolitan Subtest	29
11	Correlations Between Word Identification Skills and Comprehension	31
12	Predicted Scores on Word Identification Skills at Various Criterion Levels of Reading Comprehension	34
13	Comparisons of Skill Performances between Menominee Indian and Non-Indian Second Grade Students	37
14	Comparisons of Skill Performances between Menominee Indian and Non-Indian Fourth Grade Students	38
15	Comparisons of Skill Performances between Menominee Indian and Non-Indian Sixth Grade Students	39
16	Subjects by Grade Level and Geographic Region	45
17	Word Identification Test: Subtests and Number of Items by Grade Level and Form	48
18	Mean Scores (Raw and Percent Correct) by Grade Level	52
19	Hoyt Reliability Coefficients for Word Identification Subjects	53
20	Pearson Correlations of Word Identification Subtests with Metropolitan Subtest	55

List of Tables (continued)

Table		Page
21	Rank-ordered Correlations of Word Identification Subtests with Metropolitan Subtest	58
22	One-way Analyses of Variance Comparing Four Regions: Report of Significant F ratios ($p \leq .01$)	61
23	t-test Comparisons of Four Regions on Word Identification Test: Second Grade Level.	63
24	t-test Comparisons of Four Regions on Word Identification Test and Metropolitan Subtest: Fourth Grade Level . . .	64
25	t-test Comparisons of Four Regions on Word Identification Test and Metropolitan Subtest: Sixth Grade Level . . .	65
26	Mean Differences between the High and Low Comprehension Ability Groups, by Grade and Word Identification Subtest	67

List of Figures

Figure		Page
1	Observed levels of correlation between comprehension measure and the three word identification subskill areas tested	15
2	The predicted subskill scores for a hypothetical fourth grade subject, using linear regression equations and assuming an 80% comprehension score on the Metropolitan	35
3	Frequency distributions of scores on the Cloze Subtest and the Metropolitan Subtest for sixth grade	57

Abstract

The three studies described in this report are the first in a series of studies designed to (a) identify the subskills of word identification which correlate most highly with reading comprehension, and (b) examine various methods of assessing these subskills. The three major areas of word identification skills selected for examination were: phonics, structural analysis, and contextual analysis.

In the first study, the prototype of the Decoding (word identification) Test was administered to 282 elementary school pupils in grades one, three, and five from two Wisconsin public schools. The Reading Subtest of the Metropolitan Achievement Tests (Form F) was also administered, in order to provide information on the reading comprehension ability of each subject. The results of the first study indicated that each component of the Word Identification Test was sufficiently related to comprehension as to warrant examination in future studies.

A total of 792 students in grades two, four, and six from four Wisconsin public school districts participated in Study 2. A revised version of the Word Identification Test and the Reading Subtest of the Metropolitan Achievement Tests (Form F) were administered

to each subject. The results, as in Study 1, indicated that all three broad areas of word identification correlated significantly with comprehension. Item analyses showed good overall reliability of individual subtests. The results of Study 2 also indicated that the degree to which various word identification skills are mastered increases as children progress through the elementary grades, and that the skills which showed the highest and lowest correlations with comprehension ability were not consistent from grade to grade.

The subjects for Study 3 were 360 students in grades two, four, and six from four elementary schools. Each of the four schools represented a different geographic area of the country. As in the two previous studies, two tests were administered to each subject: a newly revised version of the Word Identification Test and the Reading Subtest of the Metropolitan Achievement Tests (Form F). Analysis of the data from Study 3 indicated that all three areas of word identification were sufficiently related to comprehension to merit inclusion in a word identification test battery. As in Studies 1 and 2, Study 3 results indicated that word identification skill mastery increases as children progress through the elementary grades. The skills showing the highest (and lowest) correlations of comprehension ability vary from grade to grade as in Study 2. The regional differences which emerged were small and appeared to reflect instructional program variations between the four regions rather than item biases within the battery.

Results from the three studies described in this report indicated that all three broad areas of word identification (phonics, structural analysis, and contextual analysis) are significantly related to comprehension. In addition, in each study, the contextual analysis subtests achieved higher correlation overall with the comprehension measure than did the structure and phonics subtests. This is most likely due to the fact that contextual analysis and reading comprehension are closely interrelated.

Changes in the subtests between Studies 2 and 3 generally resulted in higher reliability ratings for those subtests which were extensively revised. Only one newly created subtest failed to achieve an acceptable reliability level, according to results of data analysis following Study 3. The revision of subtests, however, made comparisons of results across the three studies inappropriate. These revisions account for the considerable variation of correlations between specific skill subtests and the standardized or global comprehension measure.

The ultimate goal of this research is to develop an easy-to-administer, empirically based battery of subtests for word identification skills which will provide teachers with important diagnostic information in the areas of phonics, structure, and context.

INTRODUCTION

Most elementary reading programs in the United States devote a prodigious amount of time to assessing and teaching basic word identification skills. In recent years, considerable attention has been given to defining, assessing, analyzing, and teaching the three fundamental components of word identification skills--phonic analysis, structural analysis, and contextual analysis[1] (Johnson & Pearson, 1978). Johnson and Pearson (1978) have suggested a similar approach for comprehension. Although the empirical identification of independent subskills in this area has not yet been completed, the work underway at the Center for the Study of Reading at the University of Illinois is encouraging. While progress is being made in the theoretical explorations of comprehensive skills and their applications to subskill analyses of reading instruction at the Illinois Center and elsewhere, examination of other practical and pressing issues is also warranted.

---

[1] Phonic analysis--processes which help children pronounce unfamiliar printed words as an aid to understanding their meanings. Structural analysis (morphemic analysis)--processes which help children determine the meanings of unfamiliar printed words by discerning their meaningful parts. Contextual analysis--processes which help children understand the meanings of printed words or phrases which are unfamiliar to them, or which help children learn new meanings for familiar words and phrases, by attending to the context of the material surrounding the given word or phrase.

First, how can teachers effectively assess the major components of word identification skills in order to individualize reading instruction? Because it is important for classroom teachers to assess a variety of reading subskills, what testing techniques will help teachers obtain accurate information about pupils' word identification abilities? Valid and reliable diagnostic tests that delineate areas of strength and weakness for specific word identification subskills are clearly needed.

Second, the relationship between word identification skills and comprehension abilities needs to be explored. According to Johnson and Pearson (1975):

> . . . What is at issue is whether or not mastering a specific skill improves a child's ability to read and/or comprehend running text. . . . The present question is whether or not mastery of that skill, regardless of how it is assessed, transfers to something most of us would like to call "reading".
>
> The question is, in principle, empirically solvable. Practically speaking, one probably could not determine the contribution of each individual skill. However, one could determine whether or not mastery of a set of subskills contributed to increased oral reading fluency and/or comprehension of written discourse (p. 762).

In other words, in what ways and to what degree does mastery of skills relate to reading comprehension? Do children who master word identification skills comprehend better than nonmasters? Which word identification skills relate most closely to comprehension of written discourse? Before cause-effect research studies can be undertaken, however, it seems necessary to identify potentially causative skills through the use of correlation procedures.

Third, while reading educators agree that word identification skills (often called decoding skills) are important for reading, it is not clear which correspondences, patterns, or strategies within each of the broad areas of phonics, structure, and context relate most closely to comprehension. For example, is it necessary for children to know the 61 vowel clusters in the English language and the 2 to 14 pronunciations of each one of them (e.g., ou as in soup, would, ground, sought)? Is it worth spending instructional time to learn the rule which governs the pronunciation of x according to its position within a word (i.e., xylophone, exam, tax)? Can we justify the numerous hours spent drawing short lines between syllables, because "better syllabicators are better readers"?

Hundreds of rules governing letter-sound correspondences and syllabication are taught in the schools. Most of these rules have exceptions, and many of the rules are completely erroneous. For example, the rule "when two vowels are together, the first is long and the second silent," applies only 45% of the time for words in

the primary level (Clymer, 1961) and only 18% of the time for words beyond the primary level (Emans, 1967). This rule, for example, does not apply to the words ocean, great, bread, and pause, aunt, kraut. Another rule, "when dividing VCCV words into syllables, divide between the consonants" has exceptions that affect both the syllabication and pronunciation of certain words. When syllabicating words such as summer and happen, the medial double consonants are incorrectly pronounced twice. In other words, such as father and mother, the syllables are divided after the fa- and mo-, respectively, and not between the t and h. Many rules are consistently upheld (e.g., c pronounced /s/ before e, i, and y, versus c pronounced /k/ in most other cases), but are these rules worth learning and do they warrant the instructional time?

Gray (1960), Gagne (1967), Smith (1971), and Mason (1977) have emphasized the relationships of certain word identification skills to reading comprehension. In two recent studies, the relationships between skills mastery and comprehension ability were examined, and those skills which correlated most highly with comprehension were identified.

Blohm (Note 1) conducted a pilot study to compare the comprehension test scores of children who showed mastery of the Wisconsin Design for Reading Skill Development (Design) (Otto & Askov, 1974) Word Attack skills with the comprehension scores of children who failed to show mastery. Twenty-three fourth grade "masters" (children

who scored 80% or better on 40 to 45 of the Design Word Attack skills) were compared with 23 fourth grade "nonmasters." Blohm (Note 1) found that the masters group scored two grade levels above the nonmasters group ($p < .05$) on the Metropolitan Achievement Tests. However, Blohm cautions readers of his report not to infer a cause-effect relationship.

In a similar study, Demos (1977) reported a significant positive relationship between mastery of several selected Design Word Attack skills, general reading achievement (as measured by the Gates-MacGinitie Reading Tests), and functional word attack ability (as measured by two Kurth word lists). An adapted classification scheme identified diphthongs as essential (mastered by 75% of good readers and less than 25% of poor readers), and consonant blends, short vowels, consonant digraphs, and base word endings as supportive (skills mastered by 90% of good readers and 50% or more of poor readers). No nonessential skills (skills not mastered by 75% of good readers) were identified. While Demos found that all word attack skills were essential or supportive, no single word attack skill related strongly to general reading achievement.

The reading literature indicates that children with large reading vocabularies tend to comprehend printed matter better than do children with relatively smaller reading vocabularies. We assume, therefore, that children who have learned many word identification skills have used those skills to enlarge their reading vocabularies,

and have thus become better comprehenders of what they read. It is this assumption that needs empirical evaluation and which provides the basic question for long-range investigation: "Which word identification skills and subskills correlate most highly with a standardized and widely accepted measure of reading comprehension?"

The three studies described in this paper are the first in a series of studies designed to (a) identify the subskills of word identification which appear to correlate most highly with reading comprehension, and (b) examine various methods of assessing these subskills. In the initial study, the term "decoding" was used to refer to word identification skills. In the subsequent studies, however, it was decided that the term "word identification" was a more accurate label for the broad array of skills tested.

It is intended that through this research, word identification skills which correlate most highly with reading comprehension will be identified, and that a group-administered, paper-and-pencil diagnostic test battery for assessing these skills will be developed and made available to classroom teachers.

STUDY 1

Method

Subjects

A total of 282 elementary school pupils from grades one, three, and five participated in the study. The pupils were from an elementary school in Holman, Wisconsin and a school in Waterloo, Wisconsin. Both Holman and Waterloo are rural communities with a middle class socioeconomic population. In each school, pupils from two first, third, and fifth grade classrooms participated in the study (see Table 1). The classrooms were heterogeneous in academic ability, with subjects representing a typical range of reading skills.

Stimuli

Two tests were administered to each subject: the Reading Subtest of the Metropolitan Achievement Tests (Form F), and the Decoding Test, which was constructed specifically for this study. The Decoding Test is a group-administered, paper-and-pencil test designed to assess a variety of word identification subskills within the areas of phonics, structure, and context. Three levels of the test were developed: one for first grade, one for third grade, and one for fifth grade. There was some variation among the three tests in terms of the particular skills tested and in the number of items within each subtest. A list of the subtests for each grade level, along

Table 1

Subjects by Grade Level and School District

School district	Grade			Total per school
	1	3	5	
Holman	54	44	52	150
Waterloo	42	44	46	132
Total per grade	96	88	98	282

with the number of items in each subtest is presented in Table 2. Appendix A contains a description of each of the three levels of the Decoding Test as well as sample items.

Procedure

All testing was conducted by members of the Project Staff, although first grade teachers remained in the classroom to help distribute test booklets. The Decoding Test was administered first. Upon completion of the Decoding Test, pupils were given a 20-minute break. After the break, the grade-appropriate form of the Reading Subtest of the Metropolitan Achievement Tests was administered. Total testing time, excluding the break, was approximately 1 hour to 1-1/2 hours, with the third grade subjects requiring the longest times. This was anticipated, because the third grade version of the test was longer than the other two versions.

For each subtest of the Decoding Test, the test administrator read the directions, which were printed at the top of the first page of the subtest, aloud, while the pupils read along silently. The one or two practice items at the beginning of the subtest were performed and discussed. Pupils were then instructed to work independently on the subtest. On the Phonics Subtest, however, pupils were encouraged to ask for the oral pronunciation of any unfamiliar response words, so that pupils who did not already know these words would not be at a disadvantage. Sufficient time was allowed for nearly all pupils to complete each subtest.

Table 2

Decoding Test: Subtests and Number of Items by Grade Level

Decoding Subtests	Number of items		
	Grade 1	Grade 3	Grade 5
Readiness			
Beginning Position	5	5	-
Medial Position	8	8	-
Letter Discrimination	6	6	-
Phonics			
Consonants	-	30	30
Vowels	-	22	22
Vowels and Consonants	22	-	-
Syllabication	6	6	6
Structure			
Plurals	-	5	5
Possessives	-	4	4
Affixes	-	6	6
Compound Words	-	6	6
Context			
Intruded	-	5	5
Nonsense	-	5	5
Cloze	-	6	6

The Metropolitan Subtest was administered according to the directions specified in the Teacher's Directions Manual which accompanied the packet of test booklets.

Results

Results of this pilot study are presented in Tables 3, 4, and 5; Figure 1 shows correlations across all grade levels for each subtest of the Word Identification Test with the Metropolitan Subtest. Mean scores, in both raw and percentage-correct forms, are shown for each grade level in Table 3, along with standard deviations for both the Decoding Test and the standardized Metropolitan Subtest.

Table 4 presents Pearson correlations of total Decoding Test scores with scores on the Metropolitan Subtest. These data support the notion that overall decoding skills are increasingly related to comprehension ability in successively higher grade levels, although the observed differences are relatively small.

Table 5 and Figure 1 illustrate the observed correlations between comprehension and the three broad components of the Decoding Test: structure, phonics, and context.

Although the investigation was preliminary and utilized newly created decoding tests, the relative size of the observed correlations of the phonics, structure, and context components of the Decoding Test implies that relationships exist between word identification

Table 3

Mean Scores (Raw and Percent Correct) by Grade Level

	Grade					
	1 (N=96)		3 (N=88)		5 (N=98)	
Test	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD
Decoding						
(raw)	31.28	6.19	85.66	15.54	73.28	13.60
(%)	66.5%		75.1%		77.1%	
	(47 items)		(122 items)		(103 items)	
Metropolitan Subtest						
(raw)	23.14	15.15	23.10	9.22	25.05	8.42
(%)	55.0%		51.3%		55.7%	
	(42 items)		(45 items)		(45 items)	

Table 4

Pearson Correlations Between Overall

Decoding and Comprehension, by Grade Level

Grade 1	Grade 3	Grade 5
r=.614	r=.6284	r=.6335

Note. All at $p \leq .001$.

Table 5

Correlation Matrix for Decoding Subtests and Metropolitan Subtest (Comprehension)

	Phonics	Structure	Context	Comprehension
Phonics	-	-	-	-
Structure	.559	-	-	-
Context	.651	.603	-	-
Comprehension	.591	.442	.625	-

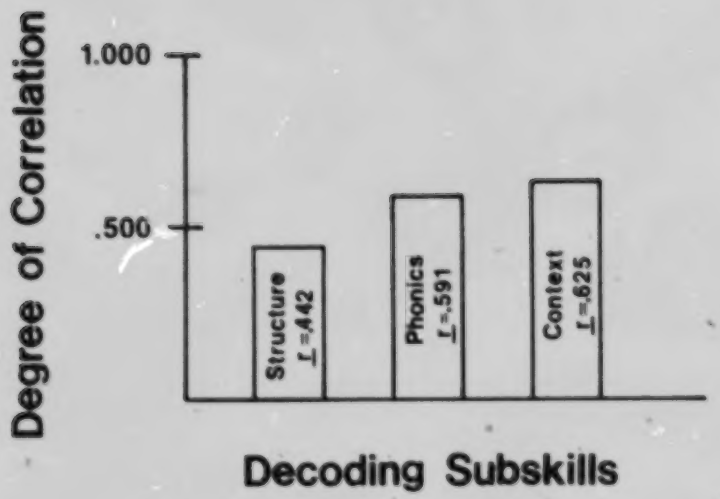


Figure 1. Observed levels of correlation between comprehension measure and the three word identification subskill areas tested.

skills and reading comprehension. Context had the highest correlation of the three word identification components, with an r of .625 ($p \leq .001$). Phonics was slightly lower, with an r of .591 ($p \leq .001$). Word structure was also correlated, but with the lowest r of .442 ($p \leq .001$).

Summary

Although this study was only a first step in the search for decoding skills which correlate highly with comprehension, some valuable insights for future studies were gained. First, the use of context must be more extensively examined, because this area achieved the highest correlation with comprehension. Efforts are being made in subsequent studies to lengthen the subtests for context and to render the context portions comparable across grade levels in order to facilitate statistical comparisons of performance.

Second, each word identification component (structure, phonics, and context) tested in this pilot study was sufficiently related to comprehension, as to warrant examination in further investigations. All the subtests were new instruments, and consequently required item analysis and revisions. For these reasons, it was decided that all three areas would be included in future studies. Moreover, within most subskills, the number of items would be increased and all subtests made as comparable as possible across grade levels.

Finally, there were no apparent patterns of subskill correlations with reading comprehension across grade levels. Because all three levels of the Decoding Test (beginning, intermediate, and advanced) varied considerably in length and content, cross-grade comparisons were deemed inappropriate. Therefore, little information was obtained concerning the possible correlations of subskill patterns with comprehension from one grade level to another. In subsequent studies, however, this issue will receive careful attention.

STUDY 2

Method

Subjects

A total of 792 students in grades two, four, and six from four Wisconsin public school districts participated in the study. One hundred and thirty-five subjects were from the Menominee Reservation School District, which is a rural area composed primarily of native American Indians of a low to middle socioeconomic population. One hundred and eighty-two subjects were from the Stoughton School District, which is a southern Wisconsin rural district composed primarily of Caucasian students from a varied socioeconomic population, ranging from lower class to upper-middle class. Three hundred and sixteen subjects were from the Brookfield School District, which is a suburb of Milwaukee composed primarily of Caucasian students from a middle to upper-middle class socioeconomic population. The final group of 159 subjects were from the Brown Deer School District. Brown Deer is also a suburban community of Milwaukee composed primarily of Caucasian students from a middle to upper-middle socioeconomic population.

The number of subjects by grade level and school district is shown in Table 6.

Table 6

Subjects by Grade Level and School District

School district	Grade 2	Grade 4	Grade 6	Total
Menominee	38	57	40	135
Stoughton	54	62	66	182
Brookfield	101	105	110	316
Brown Deer	43	41	75	159

Stimuli

Two tests were administered to each subject: the Reading Subtest of the Metropolitan Achievement Tests (Form F) and the Word Identification Test.

The Word Identification Test developed for Study 2 was a group-administered, paper-and-pencil test designed to assess a variety of word identification subskills within the areas of phonics, structure, and context. The three levels of the Decoding Test developed for Study 1 (first, third, and fifth grades) served as the basis of the development of the Word Identification Test. In Study 2, however, the test was constructed for second, fourth, and sixth grade subjects. Because Study 1 was conducted relatively late in the school year (May) and Study 2 was to be conducted relatively early in the following school year (December/January), it was believed that the same levels developed for Study 1 subjects would be appropriate for Study 2 subjects.

Two major modifications were made in the test instrument in preparation for Study 2. First, the number of items in the Structure and Context Subtests was increased and the number of items in the Phonics Subtests decreased, so that all the fourth and sixth grade subtests contained either 15, 20, 24, 25, or 26 items. Second, specific items within each subtest were revised, based on item analysis data collected in Study 1. The Living Word Vocabulary (Dale & O'Rourke, 1976) was used as a source to confirm the grade-level

appropriateness of newly developed items. Poor items were eliminated, using a criterion of below 50% in a biserial correlation as a cut-off point. In addition, the formats of individual subtests were revised. (Appendix B contains sample items from each subtest as well as a description of the modifications made on each subtest.)

In developing the Word Identification Test for Study 2, it became evident that not all subjects could take all the subtests at the appropriate grade level within a reasonable time limit. It was therefore decided that each subject would take four subtests: one subtest from each of the three main areas of Structure, Phonics, and Context, and the Compound Word Subtest. The Compound Word Subtest was administered to all subjects because the format of this subtest was new for Study 2, and was considerably different from the formats of other existing tests that had been developed to assess compound word knowledge.

Three forms, A, B, and C, of the Word Identification Test were developed at each of the three grade levels. Table 7 lists the subtests included in each form.

Procedure

Testing was conducted in the classrooms by Project Staff. Classroom teachers did not administer the tests, but in a few cases were asked to collect completed test booklets. Testing time for both the Word Identification Test and the Reading Subtest of the Metropolitan Achievement Tests ranged from 50 to 70 minutes, with each test lasting approximately 30 minutes.

Table 7

Word Identification Test: Subtests and Number of Items by Grade Level and Form

Word Identification Test	Beginning Level--Grade 2			Intermediate Level--Grade 4			Advanced Level--Grade 6		
	Form A	Form B	Form C	Form A	Form B	Form C	Form A	Form B	Form C
Structure									
Plurals (5 items)	X	X	X						
(20 items)				X			X		
Possessives (20 items)					X			X	
Affixes (20 items)						X			X
Compound Words (5 items)	X	X	X						
(15 items)				X	X	X	X	X	X
Phonics									
Consonants (20 items)	X			X			X		
Vowels (20 items)		X			X			X	
Syllabication (20 items)			X			X			X
Context									
Intruded (5 items)			X						
(25 items)				X					
(24 items)							X		
Nonsense (not given)									
(26 items)					X				
(25 items)								X	
Cloze (5 items)		X							
(25 items)						X			X
Maze (5 items)	X								

Note. The Consonants, Vowels, and Syllabication Subtests were the same across all three grades. The Plurals, Possessives, Affixes, and Compound Word Subtests were the same for fourth and sixth grades.

The first test administered was the appropriate grade level and form of the Word Identification Test. Classes within each of the three grade levels had been arbitrarily assigned to one of the three test forms. In making the assignments, an attempt was made to obtain similar numbers of subjects within each of the nine grade x format test groups.

Before beginning each subtest of the Word Identification Test, pupils were instructed to look at an item on the chalkboard. The chalkboard item was an example of the type of items that would appear on that subtest. The examiner read the chalkboard item aloud and encouraged pupils to respond. The correct answer was then wirtten on the chalkboard and discussed.

Next, the examiner read aloud the directions printed in the Word Identification Test booklets, and asked pupils to do the two or three practice items beneath the directions. After the practice items were completed and reviewed, the pupils were instructed to work independently on the subtest. On selected subtests, however, pupils were permitted to ask the examiner for assistance in reading some response items. This assistance was given so that subjects' difficulty with vocabulary would not be a confounding factor on selected subskills performances.

Time allocations for each subtest of the Word Identification Test ranged from 5 to 8 minutes. When time was up, pupils were told to stop. (A few additional minutes were provided, however,

if more than two pupils had not completed a subtest.) Then pupils were asked to turn the page in their test booklets to the next subtest.

Pupils completing a subtest before the allocated time was up were permitted to go back over their work on that subtest or on any previously administered subtest. But pupils were instructed not to begin a new subtest.

After pupils had taken all four subtests in their Word Identification Test booklets, the examiner administered the appropriate level of the Reading Subtest of the Metropolitan Achievement Tests, Form F. The Metropolitan Subtest was administered in accordance with the instructions specified in the Teacher's Directions Manual. If time remained, pupils who completed the reading subtest were encouraged to review their work on the Metropolitan Subtest.

Results

Summary statistics by grade level for Study 2 are presented in Table 8 for both the Word Identification Test and the Metropolitan Subtest.

An item analysis of the Word Identification Test was performed to identify items to be eliminated or revised in future tests. The item analysis also provided two kinds of reliability data: first, how subjects performed on individual items within a given subtest,

Table 8

Mean Scores (Raw and Percent Correct) by Grade Level

	Grade					
	2 (N=236)		4 (N=254)		6 (N=291)	
Test	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD
Word Identification						
(raw)	31.13	4.63	53.21	12.40	57.77	7.38
(%)	69.18%		66.52%		72.21%	
(range)	7-45 (of 45)		9-80 (of 80)		19-80 (of 80)	
Metropolitan Subtest						
(raw)	27.97	10.72	27.50	9.96	28.30	9.46
(%)	63.57%		61.11%		62.89%	
(range)	7-44 (of 44)		2-45 (of 45)		1-45 (of 45)	

and second, how subjects performed on individual subtests in relation to the entire Word Identification Test.

Table 9 gives Hoyt Reliability Coefficients for individual subtests correlated with the total Word Identification Test battery. The most notable exceptions to the overall good reliability of the subtests were the Maze and Cloze Subtests of the Beginning Level Context Tests. Results of these two subtests were believed to be influenced by two factors: (a) the unfamiliarity of second grade subjects with the Cloze context tests, and (b) the brevity of the subtests at the second grade level (only 5 items in each subtest). A Hoyt Reliability Coefficient of .85 is usually judged acceptable, and this level was not met by a number of subtests. Nevertheless, it was thought that because the instrument was new and would undergo additional revisions in future studies, the subtests (with the two exceptions noted above) generally performed at acceptable levels.

Table 10 shows Pearson Correlations of the various Word Identification Subtests with the Metropolitan Subtest. Other forms of analysis were not performed at this early stage, because the Word Identification Test was still in the developmental stage as an assessment device. Pearson Correlations would, however, provide at least some information about possible decoding/comprehension relationships.

On the basis of the pilot study, it was hypothesized that all three of the broad skill areas (phonics, structure, and context)

Table 9

Hoyt Reliability Coefficients for

Word Identification Subtests

Word Identification Subtests	Grade 2	Grade 4	Grade 6
Phonics			
Consonants	.77	.83	.88
Vowels	.84	.90	.85
Syllabication	.87	.85	.81
Structure			
Compound Words	.89	.81	.89
Plurals	.62	.78	.65
Possessives	-	.80	.71
Affixes	-	.81	.84
Context			
Maze	.54	-	-
Cloze	.44	.87	.90
Nonsense	-	.93	.90
Intruded	.83	.90	.87

Table 10

Pearson Correlations of Word

Identification Subtests with Metropolitan Subtest

Word Identification Subtests	Grade 2	Grade 4	Grade 6
Phonics			
Consonants	.656	.620	.567
Vowels	.653	.667	.492
Syllabication	.434	.635	.498
Structure			
Compound Words	.628	.549	.423
Plurals	.447	.562	.650
Possessives	-	.554	.597
Affixes	-	.707	.551
Context			
Maze	.720	-	-
Cloze	.480	.662	.655
Nonsense	-	.751	.774
Intruded	.786	.668	.721

would be significantly related to comprehension, in the following descending order: context, phonics, structure. The data presented in Table 10 do not offer support for this hypothesis. The range of correlations for the Structure Subtests was .423 to .707. The range for Phonic Subtests was .434 to .667. The range of Pearson Correlations for Context Subtests was .480 to .786. As in the pilot study, all subskills tested were sufficiently related to the global measure of comprehension (Metropolitan Subtest) to continue exploring such relationships in greater depth in subsequent studies.

Another interesting issue in this study was the relation of word identification subskills to comprehension across all grade levels. If crucial skills were ranked according to degree of correlation, would these rankings be consistent across grade levels? According to the results of Study 2, there were no indications of patterning. In Table 10, for example, Syllabication and Plurals for second grade subjects showed the lowest correlation with our criterion measure of comprehension. Yet, for fourth grade subjects, Compound Words and Possessives had the lowest correlations with comprehension. Correlations for sixth grade subjects indicated that Compound Words and Vowels are the subskills least related to comprehension.

A comparison of subskill tests having the highest correlations with the criterion measure yielded the same mixed results. Table 11 summarizes the observed relations between word identification and comprehension discussed above. Although Cloze showed a consistently

Table 11

Correlations Between Word Identification Skills and Comprehension

	Grade		
Correlations	2	4	6
Highest			
Phonics	Consonants (.656)	Vowels (.667)	Consonants (.567)
Structure	Compound Words (.628)	Affixes (.707)	Plurals (.680)
Context	Intruded Word format (.786)	Nonsense format (.668)	Nonsense format (.721)
Lowest			
Phonics	Syllabication (.434)	Consonants (.620)	Vowels (.492)
Structure	Plurals (.447)	Compound Words (.549)	Compound Words (.423)
Context	Cloze format (.480)	Cloze format (.662)	Cloze format (.655)

lower correlation across all three grades tested in the Context area, the correlation jumps from .480 at the second grade level to .662 at the fourth grade level and .655 at the sixth grade level. Such results suggest the following interpretations:

1. The Word Identification Test needs to be refined overall to permit more exacting comparisons across grade levels.

2. The formats which were examined within the Context area might be yielding different kinds of information about the subjects' word identification strategies and might therefore be tapping different underlying skills.

3. Children might be relying on different areas of word identification skills as their reading skills develop. That is, phonics might be relied upon most heavily in the early elementary grades. By sixth grade, however, children's strategies might be more sophisticated, utilizing, for example, contextual information for identifying unfamiliar words.

4. Critical skills for comprehension may change developmentally or remain constant across grade levels. In either case, format could have significant effects upon performance, as evidenced by the varying correlations found for the three context formats within grade levels.

Another area of interest in this study was to develop a procedure for examining the notion of skill mastery. Because of the testing design, wherein different subjects were administered different

subtests of the entire Word Identification Test, linear regression analyses were used to predict the scores on the Word Identification Subtests at a given level of performance on the Metropolitan criterion measure. Results of these analyses are summarized in Table 12 and Figure 2.

Figure 2 presents the word identification skill profile of a hypothetical fourth grade student who achieved a raw score of 80% correct on the Metropolitan Subtest. The predicted skill socres are based on the assumption that word identification skill acquisitions are linear functions of the form:

$$\underline{y}_{\underline{i}} = \underline{a}\underline{x}_{\underline{i}} + \underline{b}$$

where $\underline{x}_{\underline{i}}$ = observed score on the Metropolitan Subtest, $\underline{a}$ = slope of the line, $\underline{b}$ = intercept, and $\underline{y}$ = predicted score. Given such an assumption, and the fact that our analyses had shown that a linear function fit the data as well as or better than higher order functions, it was possible to plug in a variety of scores for comprehension. Predicted scores for various word identification skills subtests could then be calculated, generating theoretical skills profiles such as the one shown in Figure 2. In the same way, levels of mastery could be empirically, rather than arbitrarily, determined. If a student is designated a "master" with an 80% comprehension score, Figure 2 shows how his or her subskill profile might look.

Table 12

Predicted Scores on Word Identification Skills at Various Criterion Levels of Reading Comprehension

Word Identification Subtests	Grade 2			Grade 4			Grade 6		
	75%	80%	90%	75%	80%	90%	75%	80%	90%
Phonics									
Consonants	59	62	68	79	82	89	83	86	92
Vowels	56	60	67	78	82	89	87	89	94
Syllabication	62	64	68	86	88	94	88	90	94
Structure									
Compound Words	94	96	100+	64	67	74	66	69	75
Plurals	91	93	97	76	[illegible]	83	88	90	94
Possessives	-	-	-	77	79	84	82	84	89
Affixes	-	-	-	90	92	98	96	97	100+
Context									
Maze	80	84	93	-	-	-	-	-	-
Cloze	56	59	64	80	83	90	76	80	87
Nonsense	-	-	-	72	76	86	77	82	91
Intruded	80	86	98	69	73	81	72	76	83

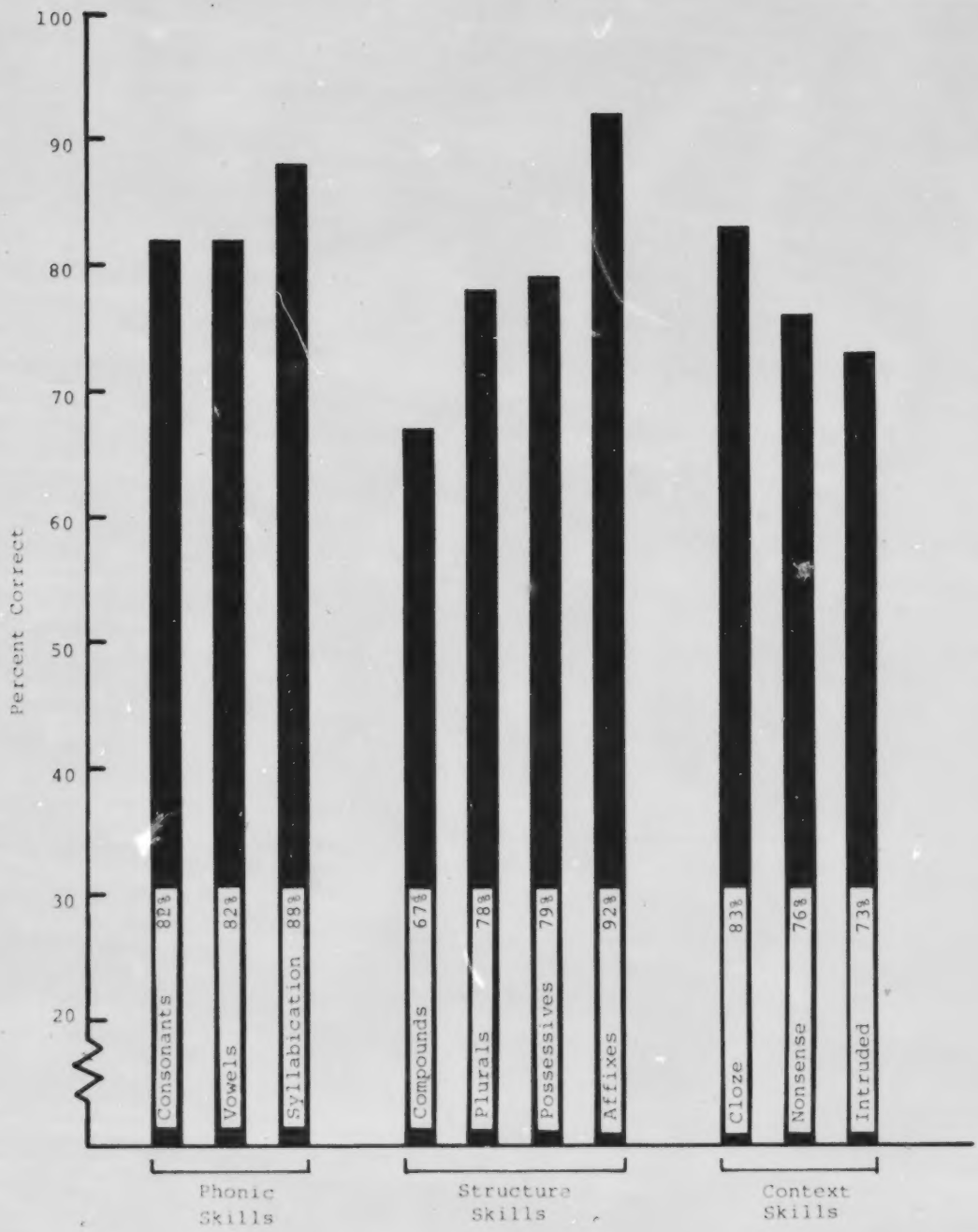


Figure 2. The predicted subskill scores for a hypothetical fourth grade subject, using linear regression equations and assuming an 80% comprehension score on the Metropolitan.

Each word identification subskill is individually related to performance on a standardized measure of reading comprehension. For example, mastery levels may range from 67% (Compound Words) to 92% (Affixes) for a fourth grade pupil who demonstrates 80% comrpehension ability. This notion of skills mastery represents a departure from traditional views, whereby mastery levels were set at one fixed value for all subskills. The present study suggests that different levels of mastery for different word identification subskills may be a more valid approach for assessing pupil performance.

A secondary objective of this study was to compare the skill profiles of Menominee Indian children with those of non-Indian children. A group of 124 (or 16%) of the subjects participating in this study were Menominee Indian children from the Menominee Reservation in central Wisconsin. Scores for these subjects were compared with scores for non-Indian subjects. t-tests were performed to discover whether or not there were differences in performance between the two groups on the various Word Identification Subtests. Results of these analyses are shown in Tables 13, 14, and 15. At all three grade levels, non-Indian subjects scored significantly higher ($p \leq .05$) on each subtest than did Menominee subjects. But when the two populations were compared, after being stratified by Metropolitan scores into high, average, and low comprehenders, the observed differences were not significant. What these data suggest then, is that an Indian child at a given level of global comprehension

Table 13

Comparisons of Skill Performances between Menominee

Indian and Non-Indian Second Grade Students

Word Identification Subtests	Raw score means		Percentage mean difference	t (absolute value)	df	Probability
	Indian	Non-Indian				
Phonics						
Consonants	6.00	10.85	24.2	5.31	79	.0001
Syllabication	5.75	12.90	35.8	6.20	77	.0001
Structure						
Compound Words	10.55	13.66	20.7	6.30	234	.0001
Plurals	3.55	4.45	18.0	4.98	234	.0001
Context						
Maze	2.00	3.52	30.4	4.78	79	.0001
Intruded	1.31	4.02	54.2	6.91	77	.0001

Table 14

Comparisons of Skill Performances between Menominee Indian and Non-Indian Fourth Grade Students

Word Identification Subtests	Raw score means		Percentage mean difference	t (absolute value)	df	Probability
	Indian	Non-Indian				
Phonics						
Consonants	11.80	14.39	13.0	2.40	84	.019
Vowels	7.25	14.90	38.2	5.34	72	.0001
Syllabication	13.07	15.97	14.6	2.58	92	.011
Structure						
Compound Words	5.17	8.91	24.9	6.71	252	.0001
Plurals	11.90	14.14	11.2	2.45	84	.016
Possessives	10.67	14.81	20.7	3.59	72	.001
Affixes	14.79	16.89	10.5	2.20	92	.030
Context						
Cloze	14.78	18.32	14.1	2.35	92	.021
Nonsense	7.42	16.27	35.4	4.20	72	.0001
Intruded	8.85	15.68	27.3	4.74	84	.0001

Table 15

Comparisons of Skill Performances between Menominee Indian and Non-Indian Sixth Grade Students

Word Identification Subtests	Raw score means		Percentage mean difference	t (absolute value)	df	Probability
	Indian	Non-Indian				
Phonics						
Consonants	9.75	16.21	32.3	6.05	95	.0001
Vowels	14.91	16.55	8.2	1.30	86	.196
Syllabication	13.61	17.12	17.6	3.84	104	.0001
Structure						
Compound Words	7.12	9.05	12.9	2.54	289	.012
Plurals	13.81	17.34	17.7	6.57	95	.0001
Possessives	12.45	15.87	17.1	3.67	86	.0001
Affixes	15.31	18.64	16.6	4.45	104	.0001
Context						
Cloze	11.92	17.24	21.3	3.11	104	.002
Nonsense	10.64	17.64	28.0	3.80	86	.0001
Intruded	7.50	17.37	39.5	9.00	95	.0001

has a skill profile similar to a non-Indian child with the same comprehension ability. But the variant background or educational program of the Indian child may delay the skill development process to some degree.

Summary

As in Study 1, this study indicated that the degree to which various word identification skills are mastered increases as children progress through the elementary grades. Also, the skills which showed the highest (and lowest) correlations with comprehension ability were not consistent from grade to grade. This finding is not surprising. Once a majority of pupils at a given grade have mastered one skill area, the correlation of that skill would be expected to have less relationship to their current comprehension ability than would more recently learned, less universally mastered, or more sophisticated skill areas.

Again, all three broad areas of word identification (contextual analysis, phonic analysis, and structural analysis), were found to have significant correlations with comprehension. Item analysis also showed good overall reliability of subtests. Revision efforts were focused on the Context Tests for Beginning Level. These were revised so they would have the same three formats as the Context Tests for the upper levels. It was decided that further testing

would be done in four regionally distinct areas to identify possible biases of the total test battery.

STUDY 3

Method

Subjects

A total of 360 students in grades two, four, and six, from four elementary schools participated in the study. Each school was in a different geographic area: North Little Rock, Arkansas; Portland, Oregon, San Gabriel, California; and Phoenix, Arizona. The schools participating in the study were selected from among several schools which had shown interest in the research being conducted by the Project.

The school in North Little Rock, Arkansas has a student population ranging from low-middle to upper-middle socioeconomic status. The students are predominantly Caucasian.

The school in Portland, Oregon is a suburban school, serving predominantly Caucasian students from a middle socioeconomic population.

The school in San Gabriel, California is a city school, serving Caucasian and Mexican-American students from a middle socioeconomic level. Non-English speaking students in the school did not participate in the study.

The school in Phoenix, Arizona is a city school serving a large population of minority students from a middle socioeconomic level.

Students in all schools were grouped heterogeneously for ability level. The number of students per grade level and geographic region is shown in Table 16.

Stimuli

Two tests were administered to each subject: the Reading Subtest of the Metropolitan Achievement Tests (Form F) and the Word Identification Test.

The Word Identification Test used in Study 3 is a group-administered, paper-and-pencil test designed to assess a variety of word identification subskills within the areas of phonics, structure, and context. The test is based on the Word Identification Test that was developed for Study 2, although five major modifications had been made in preparation for Study 3.

First, attention was given to the design of the pupil test booklet: Artwork was added to the front cover to make the booklet more attractive; an illustration of a stop-sign was placed after the last item of each subtest to indicate the end of that subtest and to remind pupils not to turn the page to a new subtest; and a STOP-page was inserted between the last subtest scheduled for Day 1 and the first subtest scheduled for Day 2.

Second, a separate Administrator's Manual was developed for each level (Beginning, Intermediate, Advanced) of the Word Identifi-

Table 16

Subjects by Grade Level and Geographic Region

Geographic region	Grade 2	Grade 4	Grade 6	Total
North Little Rock, Arkansas	19	22	25	66
Portland, Oregon	53 (2 classrooms)	39 (2 classrooms)	49 (2 classrooms)	151
San Gabriel, California	28	24	26	78
Phoenix, Arizona	19	25	21	65
			Total	360

cation Test. The Manual was to be used by classroom teachers as they administered the test. Because classroom teachers, rather than Project Staff, conducted the testing, it was necessary to have manuals with specific directions to control for consistency of test administration among classes.

Third, additional subtests were developed for the second grade level, so that there would be an appropriate form of each subtest at all grade levels. (In Study 2, second grade subjects did not receive Possessives or Affixes Subtests.)

Fourth, the number of items in four of the second grade subtests was increased from 5 to 15, in order to better assess the reliability of the instrument. The number of items for each subtest of the Context Tests for grades four and six was either increased or decreased in the hope that all Context Subtests would have 25 items. (Due to an oversight, however, the Cloze Subtest had 26 items.)

Finally, as a result of the data analysis performed for Study 2, several items on individual subtests across all three grade levels were revised. A summary of changes for each of the subtests in the Word Identification Test, as well as a sample item from each subtest is presented in Appendix C.

In both Study 2 and Study 3, three versions (forms) of the Word Identification Test were developed at each of the three grade levels. But unlike Study 2, each version in Study 3 contained all, rather than only selected, subtests within the areas of structure and phonics.

The only difference in the versions used in Study 3 was the type of the Context Subtest included. At each grade level, there were three versions of the Word Identification Test: one version containing the Cloze Subtest, one version containing the Nonsense Word Subtest, and one version containing the Intruded Word Subtest. Therefore, all second grade subjects received the same Word Identification Test, with the exception of the Context Subtest. Because the Phonics and Structure Subtests for fourth and sixth grade subjects were identical, all fourth and sixth grade subjects received the same tests, with the exception of the Context Subtests. (See Table 17 for a list of subtests of the Word Identification Test within and between grade levels.)

Procedure

During the second week in April, testing materials were mailed to the four school districts participating in the regional study, with instructions to administer the tests between April 24 and May 12, 1978. Each school district received pupil test booklets and administrator's manuals for the Word Identification Test and for the Reading Subtest of the Metropolitan Achievement Tests. A Teacher Questionnaire (see Appendix D) and general information sheets for scheduling procedures were also included.

Regional testing took place in individual classrooms over the 3-week period between April 24 and May 12, 1978. Teachers had been instructed to schedule the testing on any 3 consecutive days during

Table 17

Word Identification Test: Subtests and Number of Items by Grade Level and Form

Word Identification Test	Beginning Level--Grade 2			Intermediate Level--Grade 4			Advanced Level--Grade 6		
	Form A	Form B	Form C	Form A	Form B	Form C	Form A	Form B	Form C
Structure									
Singular and Plurals (20 items)	X	X	X	X	X	X	X	X	X
Possessives (15 items)	X	X	X						
(20 items)				X	X	X	X	X	X
Affixes (15 items)	X	X	X						
(20 items)				X	X	X	X	X	X
Compound Words (15 items)	X	X	X	X	X	X	X	X	X
Phonics									
Consonants (20 items)	X	X	X	X	X	X	X	X	X
Vowels (20 items)	X	X	X	X	X	X	X	X	X
Syllabication (20 items)	X	X	X	X	X	X	X	X	X
Context									
Intruded (15 items)			X						
(25 items)						X			X
Nonsense (15 items)	X								
(25 items)				X			X		
Cloze (15 items)		X							
(26 items)					X				
(25 items)								X	

Note. The Consonants, Vowels, Singulars and Plurals, and Syllabication Subtests were the same across all three grades. The Possessives, Affixes, and Compound Word Subtests were the same for fourth and sixth grades.

this 3-week period. The 3-day sessions, which were the same for all grades and regions, were to be scheduled as follows:

Day 1: Structure Tests (1 hour)
Singulars and Plurals
Compound Words
Possessives
Affixes

Day 2: Phonics Tests and Context Tests (1 hour)
Syllabication
Consonants
Vowels
--Break--
Context (one of three formats: Cloze, Nonsense Words, or Intruded Words)

Day 3: Reading Subtest of the Metropolitan Achievement Tests, Form F (1/2 hour)

The procedures for administering the tests were similar to those used in Study 2, except that classroom teachers, instead of the Project Staff, conducted the testing. As in Study 2, pupils were instructed to look at an item on the chalkboard before beginning each subtest of the Word Identification Test. The chalkboard item was an example of the type of items that would appear on that subtest. The teacher read the chalkboard item aloud and encouraged pupils to respond. The correct answer was then written on the chalkboard and discussed.

Following the chalkboard item, the teacher and pupils read the subtest directions printed in the Word Identification Test booklets, and performed the two or three practice items below the directions. After this practice, pupils were instructed to work independently on the subtest. On selective subtests, however, pupils were permitted to raise their hands during the testing to ask for help in reading some response items. This assistance was provided so that subjects' difficulty with vocabulary would not be a confounding factor on subskills performances.

The testing procedure was different at the second grade level on four subtests, as younger subjects required more teacher assistance. Second grade teachers were instructed to read aloud the target words of the Singulars and Plurals and Syllabication Subtests, and the response words of the Consonants and Vowels Subtests.

Time allocated for each subtest at all grade levels ranged from 8 to 15 minutes. When time was up, pupils were told to stop writing and to turn the page in their Word Identification Test booklets to the next subtest. Pupils completing a subtest before the allocated time was up were permitted to review their work on the subtest. But they were instructed not to begin a new subtest.

After pupils had taken the eight subtests in their Word Identification Test booklets over two testing sessions, the teacher administered the appropriate level of the Reading Subtest of the

Metropolitan Achievement Tests, Form F.[2] The Metropolitan Subtest was administered in accordance with directions specified in the Teacher's Directions Manuals of the Metropolitan Achievement Tests.

Results

Table 18 presents summary statistics for the three grade levels tested, for both the Word Identification Test and the Reading Subtest of the Metropolitan Achievement Tests.

An item analysis was performed on the Word Identification Test to obtain reliability data and to identify items which performed irregularly. The item analysis provided two kinds of reliability data: first, the reliability of individual test items within a given subtest, and second, the reliability of subtests in relation to the entire Word Identification Test. Table 19 presents Hoyt Reliability Coefficients for the total Word Identification Test and for individual subtests within the entire Word Identification Test. With the exception of the Possessives Subtest for second grade (Hoyt Reliability Coefficient = .49), all subtests at all

---

[2]Subjects in the Phoenix schools did not take the Reading Subtest of the Metropolitan Achievement Tests during the 3-week testing period. These pupils had taken the Metropolitan in April, and their Reading Subtest scores were made available to the Project Staff.

Table 18

Mean Scores (Raw and Percent Correct) by Grade Level

	Grade					
	2 (N=111)		4 (N=112)		6 (N=111)	
Test	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD
Word Identification						
(raw)	100.91	22.63	123.86	20.93	134.56	18.36
(%)	72.08%		77.41%		84.10%	
(range)	33-136 (of 140)		48-156 (of 160)		82-159 (of 160)	
Metropolitan Subtest						
(raw)	33.17	9.58	30.00	9.67	31.30	8.97
(%)	75.39%		66.67%		69.55%	
(range)	11-44 (of 44)		10-45 (of 45)		8-45 (of 45)	

Table 19

Hoyt Reliability Coefficients for

Word Identification Subtests

Word Identification Subtests	Grade 2	Grade 4	Grade 6
Phonics			
Consonants	.86	.83	.86
Vowels	.85	.84	.86
Syllabication	.88	.80	.79
Structure			
Compound Words	.88	.71	.78
Singulars and Plurals	.81	.81	.76
Possessives	.49	.74	.70
Affixes	.78	.75	.63
Context			
Cloze	.93	.82	.92
Nonsense Word	.87	.90	.87
Intruded Word	.79	.83	.87
Total Word Identification Test	.96	.95	.95

three levels of the Word Identification Test achieved high reliability ratings. An item analysis of the Possessives Subtest did not identify any poor items; the low reliability coefficient for this subtest was probably due to the fact that this subtest had only two response choices ("yes" or "no"). Another factor may be that possessives are not directly taught as a part of most second grade curricula. Children's responses would probably be based on guesses, and would therefore approximate random chance in this 15-item, "Yes-No" format.

A central focus of this investigation has been to identify possible relationships between word identification skills which are commonly taught and tested, and performance on a global measure of comprehension (the Metropolitan Subtest). Results of Studies 1 and 2, described earlier in this report, had indicated that not only were the three broad skill areas (phonics, structure, and context) significantly related to comprehension, but also that each individual subskill was likewise related to comprehension. Table 20 shows Pearson correlations of the various Word Identification Subtests with the Reading Subtest of the Metropolitan Achievement Tests. These findings essentially replicate and confirm the findings of the first two studies. The range of correlations for the Phonics Subtests was .436 to .646, the range of correlations for the Structure Subtests was .311 to .767, and the range of correlations for the Context Subtests was .048 to .895. Of particular note is the exceptionally low correlation (.048) for the sixth grade level of the Cloze Subtest with the Metropolitan Subtest. Yet, the Cloze Subtest had a

Table 20

Pearson Correlations of Word

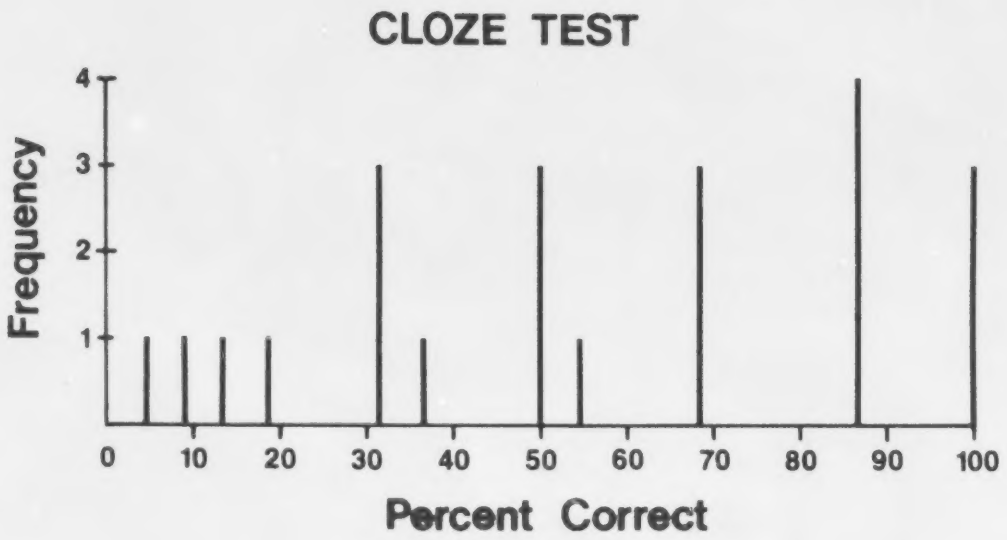
Identification Subtests with Metropolitan Subtest

Word Identification Subtests	Grade 2	Grade 4	Grade 6
Phonics			
Consonants	.646	.489	.501
Vowels	.627	.563	.546
Syllabication	.529	.436	.504
Structure			
Compound Words	.530	.311	.566
Singulars and Plurals	.358	.566	.597
Possessives	.458	.490	.534
Affixes	.767	.650	.586
Context			
Cloze	.841	.521	.048
Nonsense Word	.786	.775	.681
Intruded Word	.742	.895	.593
Total Word Identification Test	.777	.720	.734

Hoyt Reliability Coefficient of .92. The item analysis indicated that of the 25 items on the subtest, only 5 were missed by more than 6 of the 22 subjects who took the test. Clearly, there was a marked ceiling effect on this subtest. Also, as shown in Figure 3, the frequency distributions of scores on the Cloze Subtest and the Metropolitan Subtest shared little overlapping area.

Another question of interest in this study, as in Study 2, was the relationship of word identification subskills to comprehension across grade levels. If crucial word identification skills were ranked by degree of correlation within grades, would these rankings be consistent across grades? The rank-ordered correlations, from highest to lowest, of the 10 subskills with comprehension are presented in Table 21. At first glance, no striking pattern is evident. For example, of the 10 subtests, Cloze ranks first at the second grade level, sixth at the fourth grade level, and tenth at the sixth grade level. A closer examination of Table 21, however, suggests that certain subtests' rankings are quite stable across the grade levels tested. The Nonsense Word Subtest ranked either first or second for each of the three grade levels. Affixes ranked third, third, and fourth for grades two, four, and six, respectively. Vowels ranked sixth, fifth, and sixth. Syllabication achieved the low but stable rankings of eighth, ninth, and eighth.

It seems apparent that a pattern of format consistency across grade levels can have useful implications for assessment and instruc-



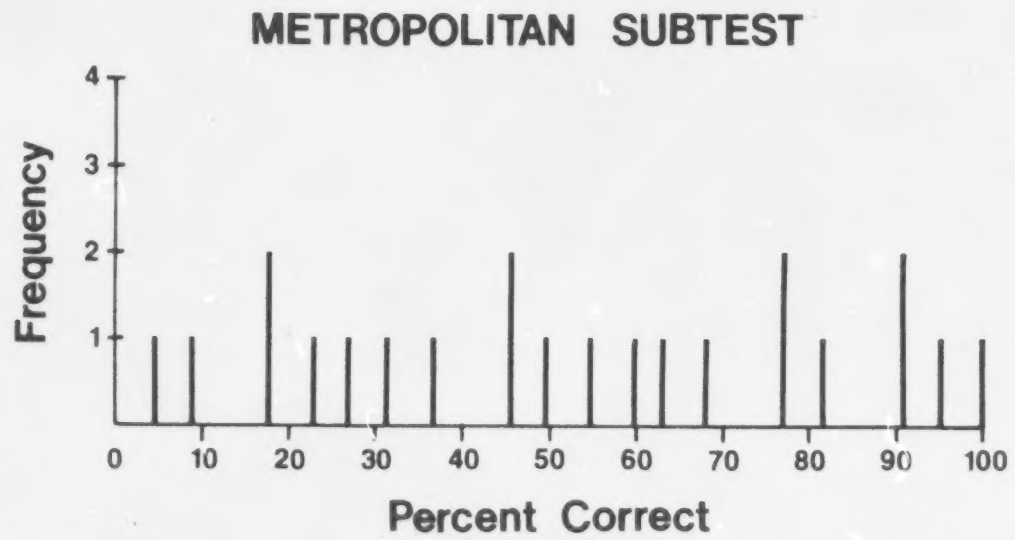


Figure 3. Frequency distributions of scores on the Cloze Subtest and the Metropolitan Subtest for sixth grade.

Table 21

Rank-ordered Correlations of Word Identification Subtests with Metropolitan Subtest

	Grade		
Ranking	2	4	6
1	Cloze (.841)	Intruded Word (.895)	Nonsense Word (.681)
2	Nonsense Word (.786)	Nonsense Word (.775)	Singulars and Plurals (.597)
3	Affixes (.767)	Affixes (.650)	Intruded Word (.593)
4	Intruded Word (.742)	Singulars and Plurals (.566)	Affixes (.586)
5	Consonants (.646)	Vowels (.563)	Compound Words (.566)
6	Vowels (.627)	Cloze (.521)	Vowels (.546)
7	Compound Words (.530)	Possessives (.490)	Possessives (.534)
8	Syllabication (.529)	Consonants (.489)	Syllabication (.504)
9	Possessives (.458)	Syllabication (.436)	Consonants (.501)
10	Singulars and Plurals (.358)	Compound Words (.311)	Cloze (.048)

tion of particular word identification skills. Although all three context formats achieved high reliability (see Table 19) at the grade levels tested, only the Nonsense Word format had a consistently high correlation with the Metropolitan Subtest across grades. Within the broad area of structure, once again the Affixes, Compound Words, Possessives, and Singulars and Plurals Subtests were comparably reliable. Yet only the Affixes Subtest achieved high and consistent correlations with comprehension. The other subskill areas either ranked lower overall, or had widely fluctuating degrees of correlation with comprehension (e.g., Singulars and Plurals had rankings of 10, 4, and 2 across the grades tested). To some degree, interpretation of these results must take into account the sequence in which phonic and structural skills are introduced and taught. If instruction in singulars and plurals is not emphasized until the middle elementary grades, then second grade students' performance on subtests assessing this skill will not show a relationship to comprehension ability regardless of the reliability of the assessment method.

Another consideration in interpreting the results of this analysis is the task similarity of the highly and consistently ranked subskills to the tasks important to global comprehension itself. For example, the task required for the Nonsense Word format of the Context Test (see Appendix C) could be regarded as similar to the task required for the Cloze format--in both formats

a student has to provide a word in a sentence. In the Cloze format there is a blank space, while in the Nonsense Word format there is a synthetic word. The Cloze format is a well-known measure for assessing comprehension. It is not surprising, therefore, that students who achieve high scores on the Cloze and Nonsense Word formats, also achieve high scores on the Metropolitan Subtest. The Affixes Subtest, which achieved a fairly high ranking across grades for its association with comprehension, also utilized a Cloze-type format (see Appendix C).

Vowels ranked slightly higher than consonants across grades. This was not unexpected because vowel complexity is known to be greater than consonant complexity; therefore, vowels are less likely to be mastered than consonants--even by the end of sixth grade. As indicated in Table 21, the Consonants Subtest had successively lower correlation with the Metropolitan Subtest as grade level increased. This is probalby due to the ceiling effect that was apparent on the Consonants Subtest.

One of the major purposes of this study was to examine the Word Identification data for regional differences. The four regions tested were compared using a one-way analysis of variance technique and by t-tests between the four testing locations.

Table 22 lists only those F-tests which were significant with probability $\leq .01$. In the analysis of variance tests, only the second and sixth grade levels of the Word Identification Test

Table 22

One-way Analyses of Variance Comparing Four Regions:

Report of Significant F ratios ($p \leq .01$)

Word Identification Subtests	Grade 2		4	6	
Consonants	$F = 6.10$	$p = .007$	---	---	---
Vowels	$F = 5.15$	$p = .0023$	---	$F = 4.47$	$p = .0053$
Syllabication	$F = 8.93$	$p = .0000$	---	$F = 5.79$	$p = .0011$
Singulars and Plurals	---	---	---	$F = 4.37$	$p = .0061$
Possessives	---	---	---	$F = 6.21$	$p = .0006$
Compound Words	---	---	---	$F = 4.21$	$p = .0074$
Total Word Identification Test	$F = 8.71$	$p = .0000$	---	$F = 5.27$	$p = .0020$

showed regional differences. The regional groups were compared via t-tests to determine where the regional differences appear within the Word Identification battery. These comparisons are summarized in detail in Tables 23, 24, and 25 for grades two, four, and six, respectively. The three Context formats are not included in Tables 22 through 25. Comparisons could not be made between regions on this skill area because each region received only one of the three formats. In general, the differences which appeared on the other individual subtests were between Portland and the other three regions. Except for fourth grade, all the significant differences in subtest performance were between Portland and the other three regions. At the fourth grade level, three significant comparisons involved Portland, while the rest involved Phoenix and Little Rock (four comparisons), and San Gabriel and Phoenix (one comparison).

Overall, the regional differences were small, with t-values ranging from 2.44 to 4.85. Although care was taken in selecting the regional testing locations to ensure a reasonable degree of comparability, it was apparent from these analyses that the Portland subjects achieved lower scores on all measures than did subjects from the other three locations. For this reason, it was decided that certain planned analyses involving stratifications of the total sample into comprehension ability subgroups would not contribute useful information. It had been hoped that the data set would display sufficient homogeneity between regions so that

Table 23

t-test Comparisons of Four Regions on Word Identification

Test: Second Grade Level

Comparison	Subtest	t-value	Probability
San Gabriel vs. Portland	Consonants	3.51	.000
Phoenix vs. Portland	Consonants	3.28	.001
Little Rock vs. Portland	Consonants	2.49	.014
San Gabriel vs. Portland	Vowels	3.42	.000
Phoenix vs. Portland	Vowels	2.85	.005
San Gabriel vs. Portland	Syllabication	3.82	.000
Phoenix vs. Portland	Syllabication	3.98	.000
Little Rock vs. Portland	Syllabication	3.63	.000
San Gabriel vs. Portland	Singulars and Plurals	2.56	.012
Phoenix vs. Portland	Singulars and Plurals	2.79	.006
Phoenix vs. Portland	Possessives	2.54	.013
Phoenix vs. Portland	Affixes	2.65	.009
Little Rock vs. Portland	Affixes	2.59	.011
San Gabriel vs. Portland	Total Word Identification Test	3.88	.000
Phoenix vs. Portland	Total Word Identification Test	3.96	.000
Little Rock vs. Portland	Total Word Identification Test	3.43	.000

Table 24

t-test Comparisons of Four Regions on Word Identification Test and Metropolitan Subtest: Fourth Grade Level

Comparison	Subtest	t-value	Probability
Phoenix vs. Little Rock	Singulars and Plurals	2.53	.013
Little Rock vs. Phoenix	Singulars and Plurals	2.79	.006
Phoenix vs. Little Rock	Possessives	3.05	.003
Little Rock vs. Portland	Possessives	2.58	.011
Little Rock vs. Portland	Affixes	2.49	.014
San Gabriel vs. Phoenix	Metropolitan Subtest	2.60	.011
Phoenix vs. Portland	Metropolitan Subtest	3.53	.001
Phoenix vs. Little Rock	Metropolitan Subtest	4.85	.000

Table 25

t-test Comparisons of Four Regions on Word Identification Test and Metropolitan Subtest: Sixth Grade Level

Comparison	Subtest	*t*-value	Probability
Phoenix vs. Portland	Vowels	2.62	.010
Little Rock vs. Portland	Vowels	3.21	.002
San Gabriel vs. Portland	Syllabication	3.01	.003
Phoenix vs. Portland	Syllabication	2.47	.015
Little Rock vs. Portland	Syllabication	3.61	.000
Little Rock vs. Portland	Singulars and Plurals	3.62	.000
San Gabriel vs. Portland	Possessives	2.65	.009
Little Rock vs. Portland	Possessives	4.05	.000
San Gabriel vs. Portland	Compound Words	2.44	.016
San Gabriel vs. Portland	Total Word Identification Test	2.48	.015
Little Rock vs. Portland	Total Word Identification Test	3.64	.000
Little Rock vs. Portland	Metropolitan Subtest	2.86	.005

it could be used in examining possible differences in performance by same comprehension-ability groupings on the three context formats.

One stratification procedure was performed to see whether subjects with low comprehension scores on the Metropolitan Subtest (below 23 of 44 or 45 possible) were indeed performing lower on each of the Word Identification Subtests than were subjects whose Metropolitan scores were high (above 31 out of a possible score of 44 or 45). Table 26 presents the differences in the mean raw scores of the high and low groups on each of the Word Identification Subtests. Both the overall F-tests and t-tests which were performed on the two groups indicated that the high comprehension group did much better on every subtest than did the low comprehension level group. The t-values were large and significant beyond the .00001 probability level.

Teacher Questionnaires had been distributed to all teachers whose students had participated in the regional study (see Appendix D). The questionnaires were designed to provide information in two areas. First, each teacher was asked to list the names of the children who were the five best readers in the class and then, the five children who had the most difficulty with reading. Teachers were also asked to list the names of any children who had been identified for special education services. Accuracy ratios were then determined for each teacher by computing the number of students which the

Table 26

Mean Differences between the High and Low Comprehension Ability Groups, by Grade and Word Identification Subtest

Word Identification Subtests	Number of items	Grade 2	Number of items	Grade 4	Number of items	Grade 6
Phonics						
Consonants	20	7.09	20	4.27	20	4.31
Vowels	20	6.74	20	4.98	20	4.98
Syllabication	20	6.11	20	2.83	20	5.62
Structure						
Compound Words	15	2.69	15	1.92	15	4.10
Singulars and Plurals	20	3.31	20	4.57	20	4.46
Possessives	15	2.18	20	3.58	20	3.02
Affixes	15	5.54	20	3.81	20	2.18
Context (for all three formats)	15	7.21	25	6.64	25	6.27
Total Word Identification Test	140	40.86	160	32.6	160	32.7

teacher ranked in the top or bottom five, and by comparing this list to a list of students ranking in the top and bottom five based on scores achieved on the Word Identification Test. For example, if three of the top five readers, as judged by their teacher, also scored in the top five on the Word Identification Test, the teacher's accuracy score was three.

This comparison procedure yielded the following results: 64% of the top-scoring students on the Word Identification Test were rated by their teachers as being among the top five readers in their class. Seventy-six percent of the low-scoring students were rated by their teachers as being among the poorest five readers in the class. Apparently, teachers can more readily identify the poorest readers than the best readers, at least as far as those reading abilities are reflected in the Word Identification Test used in this study, though overall teachers did well with predicting both.

Summary

All three broad areas of the test battery (phonics, structure, and context) obtained high overall reliability ratings. Only the Possessives Subtest for Beginning Level achieved a low rating, with a Hoyt Reliability Coefficient of .49. This subtest will be completely revised and will be incorporated into a test of Inflected Endings in future testing sessions. It is clear that the revision

of the Beginning Level Context Tests for Study 3 had a positive effect in increasing the reliability ratings of these subtests. The Hoyt Reliability Ratings for these subtests were .93, .89, and .79 as compared to the low (.44, .54, and .83) reliability ratings in Study 2. A marked ceiling effect on the Cloze Subtest of the Advanced Level Context Tests was noted.

The cross-grade comparisons of the 10 subskill areas tested in Study 3 did yield several consistent patterns: Vowels and Syllabication achieved relatively low but stable correlations across grades. The Nonsense Word format within the Context area achieved high rankings across grades. Affixes achieved high and consistent correlations within the Structural Analysis area.

The regional differences which emerged from Study 3 appeared between Portland, Oregon and the other three regional test sites. Generally, regional differences were small and appeared to reflect instructional program variations between the four regions rather than biases of any particular items within the battery.

The subskills tested in Study 3 again showed that the three broad skill areas (phonics, structure, and context) were sufficiently related to comprehension to merit inclusion in a word identification test battery. As in Studies 1 and 2, Study 3 results indicated that word identification skill mastery increases as children progress through the elementary grades. The skills showing the highest (and lowest) correlations with comprehension ability varied from grade to

grade as in Study 2. Furthermore, only 50% of the subtests showing the highest correlation with comprehension within a subskill area in Study 2 (see Table 11) also correlated highest with comprehension on Study 3. Conclusions could not be drawn from this disagreement between Studies 2 and 3 because test revisions were a confounding factor in comparing results of the two studies.

CONCLUSION

The research described in this report is concerned with three facets of word identification assessment: (a) the relationship of commonly taught word identification subskills to global comprehension ability; (b) the selection of appropriate test formats and target items for subskill assessment; and (c) the development of procedures for administrating the test battery.

Results from the three studies described in this report indicate that all three broad areas of word identification (phonics, structure, and context) are significantly related to comprehension. In addition, in each study, the Context Subtests achieved higher correlation overall with the comprehension measure than did the Structure and phonics Subtests. This is most likely due to the fact that contextual analysis and reading comprehension are so closely interrelated.

Changes in subtests between Studies 2 and 3 generally resulted in higher reliability ratings for those subtests which were extensively revised. Results of data analysis following Study 3 showed that only one newly created subtest (Possessives at the Beginning Level) failed to achieve an acceptable reliability level. The revisions of subtests, however, made comparisons of results across

the three studies inappropriate. These resivions explain why the relative degree of correlation between specific skill subtests and the comprehension measure varied considerably across the three studies.

These studies were preliminary, and were intended as much to aid in the refinement of the assessment instrument as to answer questions about the relationship of word identification skills to comprehension ability. The results suggest the following questions for consideration in future studies:

1. If elementary school children use word identification (decoding) skills to enlarge their reading vocabularies, which, in turn, raises their comprehension levels, do they use these skills more adeptly as they progress through the grades? One goal of future studies will be to identify more specific elements of word identification that correlate most highly with reading comprehension.

2. Could skill profiles be generated for specific word identification skills at various grade levels which might show the relative contribution of each of these skills to comprehension over time?

3. Could future testing strategies be designed to examine more closely the notion of mastery, so that for each subskill, a meaningful criterion could be empirically determined which would differentiate between "good" and "poor" comprehenders? This kind

of information would have the greatest implications in classrooms, where instructional objectives and apportioning instructional time are primary considerations for effective and efficient teaching.

Through extensive data collection and analysis of student performance on the subtests developed to date, a reliable and easily administered battery of word identification subskills has been created. Future studies will refine measurement procedures and instruments, and examine more precisely the different components of word identification skills. For the next year, research efforts will be directed toward one subtest at a time for the purposes of:

1. developing an empirically based rationale for the selection of target items; and

2. establishing concurrent validity for each subtest by the administration of oral reading passages with embedded target items for a randomly selected subset of each test sample.

The phonics subtests are currently being revised using Venezky's tabulations of spelling-to-sound correspondence frequencies as a basis for selecting target items (Venezky, 1970, Note 2). Distractors in the Phonics Test are being created according to a documented rationale, using confusion matrix studies of auditory and visual perceptions of letter sounds and features as a foundation for response choices.

A substudy is also being planned to examine the performances of elementary school children on two commonly used formats of

phonic assessment: (a) a decoding task which will involve the pronunciation of sounds from letter strings; and (b) an encoding task which will involve the spelling of stimuli presented aurally. This substudy will focus on those phonic correspondences which are not bidirectional, in an effort to determine which assessment mode, decoding or encoding, is most appropriate for phonics.

The subtests in the structural analysis area are also being revised. Two subskills which had been previously assessed by separate subtests (Possessives and Singulars and Plurals) will now be incorporated into a single subtest of Inflected Endings. Target items for both the Inflected Endings Subtest and the Affixes Subtest will be selected according to frequency data and grade level appropriateness obtained from a variety of sources. The Compound Words Subtest will be examined, and a revised version will be administered to a sample of fourth and fifth grade subjects.

The third area of the Word Identification Test, Contextual Analysis, is also being re-examined. All three studies have indicated that no one format for assessing context is best for all three grade levels. The three context formats (cloze, nonsense word, and intruded word) will therefore be revised, and grade-appropriate forms administered to samples of second, fourth, and sixth grade subjects. Research will look at types of contextual clues readers use rather than examining methods of testing contextual usage.

Each subtest in the Word Identification Test will initially be composed of a pool of items (e.g., 171 items for the Phonics Subtest) which will later be reduced to a smaller, but predictive subset (perhaps 40 items for the Phonics Subtest).

The major area of concern in these future substudies will be first, the development of empirically based item pools, and then later, the reduction to workable subsets within each word identification skill area. A global measure of comprehension will be used only after the final test battery is ready for administration. Real reading tasks (oral) will be given to subjects involved in the pilot tests of each subtest to establish concurrent or face validity for the various subtests.

The ultimate goal of this research is to develop an easy-to-administer, empirically based battery of subtests for word identification skills, which will provide teachers with important diagnostic information in the areas of phonics, structure, and context.

Reference Notes

1. Blohm, P. Wisconsin design word attack skills and overall reading achievement. Unpublished manuscript, University of Wisconsin-Madison, 1975.

2. Venezky, R. L. A computer program for deriving spelling to sound correlations. Master's thesis, Cornell University, 1962. Published in part in H. Levin, et al. A basic research program on reading. Ithaca, NY: Cornell University, 1963.

References

Clymer, T. The utility of phonic generalizations in the primary grades. The Reading Teacher, 1963, 16, 252-258.

Dale, E., & O'Rourke, J. The living word vocabulary. Elgin, IL: Dome, Inc., 1976.

Demos, E. S. The relationship between mastery of selected word attack skills and reading performance. Unpublished doctoral dissertation, University of Wisconsin-Madison, 1977.

Emans, R. The usefulness of phonic generalizations above the primary level. The Reading Teacher, 1967, 20, 419-425.

Gagne, R. M. The conditions of learning (2nd ed.). New York: Appleton-Century-Crofts, 1967.

Gray, W. S. Reading and physiology and psychology of reading. In C. W. Harris (Ed.), Encyclopedia of educational research. New York: Macmillan, 1960.

Johnson, D. D., & Pearson, P. D. Skills management systems: A critique. The Reading Teacher, 1975, 28(8), 757-764.

Johnson, D. D. & Pearson, P. D. Teaching reading vocabulary. New York: Holt, Rinehart, & Winston, 1978.

Mason, J. M. Questioning the notion of independent processing stages in reading. Journal of Educational Psychology, 1977, 69(3), 288-297.

Otto, W., & Askov, E. Wisconsin design for reading skill development: Rationale and guidelines. Minneapolis, MN: National Computer Systems, 1974.

Pearson, P. D., & Johnson, D. D. Teaching reading comprehension. New York: Holt, Rinehart, & Winston, 1978.

Smith, F. Understanding reading. New York: Holt, Rinehart, & Winston, 1971.

Venezky, R. L. The structure of English orthography. The Hague: Mouton, 1970.

Appendix A

Sample Items from Each Subtest of the Decoding Test, Spring 1977

A Description of the Decoding Test by Grade Level

Spring 1977

First Grade Test

The First Grade Test assessed phonics, but did not include subtests on structure and context. The test also included a set of subtests assessing reading readiness.

Reading Readiness

Beginning position. Two words were read aloud by the test administrator and subjects had to decide whether or not the two words began with the same sound. Subjects indicated their response by circling either Yes or No.

Medial position. Two words were read aloud by the test administrator and subjects had to decide whether or not the two words had the same medial sound. Subjects indicated their response by circling either Yes or No.

Letter discrimination. Subjects had to look at a row containing four pairs of letters and circle the pair of letters which was different from the rest of the pairs in the row.

Phonics

Vowels and Consonants. Subjects were directed to look at a synthetic word and then at three real response words, each having an underlined part. Subjects had to circle the word in the row whose underlined part sounded the same as the underlined part of the synthetic word.

Syllabication. Words were read aloud by the test administrator and subjects had to circle the number of syllables in each word. The numerals ranged from one to five.

Third Grade Test

The version of the test for third grade subjects included subtests in all three areas of decoding: phonics, structure, and context. In addition, this version of the test included the subtests on reading readiness.

Reading Readiness

The three subtests, beginning position, medial position, and letter discrimination were the same as in the First Grade Test. These Reading Readiness subtests were included in the Third Grade Test only to lend further insight into the evaluation of first grade subjects' performance on these subtests.

Phonics

Consonants. This subtest was of the same format and included many of the same items as the Vowels and Consonants subtest used in the First Grade Test. Unlike the combined Vowels and Consonants subtest in the First Grade Test, all the underlined letter(s) in the synthetic words were consonants.

Vowels. This subtest was also of the same format and included many of the same items as the Vowels and Consonants subtest used with first grade subjects in Study 1. All the underlined letter(s) were vowels.

Syllabication. This subtest was the same as the Syllabication subtest in the First Grade Test.

Structure

Plurals. Subjects had to read each target word and decide if the word described one thing or more than one thing. The subjects indicated their response by circling either ONE or MORE.

Possessives. Subjects had to read a sentence that had the target word underlined and decide if the target word showed belonging (ownership). If it did, the subject circled YES; if it did not, the subject circled NO.

Affixes. Subjects had to read a sentence with a blank (indicating that a word was missing) and select the word, from among three alternatives, that best completed the sentence. The three foils all contained the same root word but varied in the prefix or suffix.

Compound Words. Subjects had to look at a synthetic compound word which was composed of two real words and draw a dividing line between the two real words.

Context

Intruded Word. Subjects had to read a sentence in which an extra word had been inserted and then identify the intruded word by circling it.

Nonsense Word. Subjects had to read a sentence in which a nonsense word had been substituted for a real word. Subjects then had to select, from the three alternatives below the sentence, the word which best replaced the nonsense word.

Cloze. Subjects had to read a sentence in which a blank line had been substituted in place of a real word and then print a word in the blank that completed the sentence.

Fifth Grade Test

The Fifth Grade Test was the same as the Third Grade Test, only it did not include the Reading Readiness Subtests.

Appendix B

Sample Items from Each Subtest of the Word Identification Test,

Winter 1977-78

A Description of the Word Identification Test

Winter 1977-78

Reading Readiness

The Reading Readiness Subtests, which were given to first and third grade subjects in Study 1, were not included in the Word Identification Test developed for Study 2. The data collected in Study 1 indicated that the first grade subjects had already acquired the skills assessed in these subtests.

Phonics

Vowels and Consonants. This subtest, which was administered to first grade subjects in Study 1, was eliminated. It was inferred from the data that separate subtests for vowels and consonants would be more effective and efficient than a combined test. The same Phonics Subtests were used for all three grade levels of the Word Identification Test.

Vowels. The format remained the same, although some minor revisions were made in the wording of the directions printed at the top of the page. One item was revised as a result of the item analysis. The total number of items in the subtest was reduced from 30 to 20.

Consonants. The format remained the same, although some minor revisions were made in the wording of the directions. One practice item and one test item were revised. The total number of items in the subtest was reduced from 30 to 20.

Syllabication. The format remained the same, but the examiner did not read the target words aloud, as was done in Study 1. A second practice item was added, and the number of test items was increased from 6 to 20. Only two of the items from the original version of the test were used in Study 2.

Structure

Plurals. The test developed for second grade subjects was exactly the same as the test that had been used in Study 1. Several modifications were made, however, in the test developed for fourth and sixth grade subjects. First, a third response choice (NEITHER) was added. Second, a third practice item was added to illustrate the use of the new response choice. Finally, the number of test items was increased from 5 to 20, retaining only 2 of the items from the original version of the test.

Possessives. The format remained the same, but the number of items was increased from 4 to 20. All of the items from the original test were used in the new version, although one of the items was revised slightly.

Affixes. The format remained the same, but the number of items was increased from 6 to 20. Five of the items were from the original test.

Compound Words. The format developed for second grade subjects was the same as the format used in Study 1. The number of items was increased from 5 to 15, and a practice item was added.

A new format was developed for fourth and sixth grade subjects. Subjects were required to examine 15 real compound words and to classify each compound word by writing 1, 2, or 3: 1 represented compound words which could be defined using the phrase of a; 2 represented compound words which could be defined using the phrase for a; and 3 represented compound words which could be defined using the phrase like a.

Context

Three formats of Context Tests were used in Study 1: Intruded, Cloze, and Nonsense. All three of the formats were again used with third

and fifth grade subjects in Study 2. Second grade subjects received Intruded and Cloze, but were given a Maze format in place of the Nonsense format. (First grade subjects in Study 1 did not receive any Context Tests.)

All three formats of the Context Tests developed for second grade subjects in Study 2 were based on a variation of the five sentences which had comprised the Nonsense format in Study 1. The same basic sentence was used for all three subtests in order to make the subtests as comparable as possible. (In Study 1, each format had been based on a unique set of sentences, so comparisons were unreliable.)

The Context Subtests developed for fourth and sixth grade subjects in Study 2 were of the same three formats that were used in Study 1. Instead of using independent sentences, however each test contained a passage taken from the appropriate level of the Ginn 720 basal reading series (Level 8 and Level 11). The same two passages (a fourth grade passage and a sixth grade passage) were then adapted for all three types of Context Tests (Intruded, Cloze, and Nonsense). Each passage contained 24 to 26 items.

Directions for all three levels (second, fourth, and sixth) of the Context Subtests were expanded from what had been used in Study 1; a third practice item was added to the fourth and sixth grade subtests.

Sample Items from Each Subtest of

the Decoding Test

Spring 1977

Reading Readiness: Beginning Position (First Grade Test, Third Grade Test)

(Teacher read "fat....sat.")	1.	YES	NO
(Teacher read "mat....met.")	2.	YES	NO

Reading Readiness: Medial Position (First Grade Test, Third Grade Test)

(Teacher read "kit....fat.")	1.	YES	NO
(Teacher read "pop....dot.")	2.	YES	NO

Reading Readiness: Letter Discrimination (First Grade Test, Third Grade Test)

1.	ru	ru	rv	ru
2.	go	yo	go	go

Phonics: Vowels and Consonants (First Grade Test)

1.	zell	feel	real	belted
2.	fon	oat	din	mar

Phonics: Consonants (Third Grade Test, Fifth Grade Test)

1.	yed	saw	you	ned
2.	hinny	knew	gong	shimmy

Phonics: Vowels (Third Grade Test, Fifth Grade Test)

1.	zell	feel	real	belted
2.	crume	fur	umbrella	room

Phonics: Syllabication (First Grade Test, Third Grade Test, Fifth Grade Test)

1.	company	1	2	3	4	5
2.	whose	1	2	3	4	5

Structure: Plurals (Third Grade Test, Fifth Grade Test)

1.	cows	ONE	MORE
2.	kiss	ONE	MORE

Structure: Possessives (Third Grade Test, Fifth Grade Test)

1.	Sam's a good dog.	YES	NO
2.	The twins' clothes were just alike.	YES	NO

Structure: Affixes (Third Grade Test, Fifth Grade Test)

1.	Of all the pictures, hers turned out the _____.		
	nicer	nicest	nicely
2.	Blind people are _____.		
	sightless	sighted	sightful

Structure: Compound Words (Third Grade Test, Fifth Grade Test)

1.	r a m p w a l k
2.	b o a t d r o p

Context: Intruded Word (Third Grade Test, Fifth Grade Test)

1.	Jim wanted his mother six to meet his new friends.
2.	The car started down light the city street.

Context: Nonsense Word (Third Grade Test, Fifth Grade Test)

1.	The children splashed and played on the raft at the city <u>flugle</u>. hall beach dump
2.	<u>Zats</u> cry when they are wet or hungry or sick because they do not talk yet. trees babies dogs

Context: Cloze (Third Grade Test, Fifth Grade Test)

1.	Fred never liked to go to _____ at night, even when he was very tired.
2.	Everyone had fun at the birthday _____.

Sample Items from Each Subtest of the

Word Identification Test

Winter 1977-78

Phonics: Consonants (Second Grade Test, Fourth Grade Test, Sixth Grade Test)

1.	yed	saw	you	ned
2.	hinny	gong	shimmy	knew

Phonics: Vowels (Second Grade Test, Fourth Grade Test, Sixth Grade Test)

1.	zell	feel	real	belted
2.	crube	fur	umbrella	room

Phonics: Syllabication (Second Grade Test, Fourth Grade Test, Sixth Grade Test)

1.	American	1	2	3	4	5
2.	called	1	2	3	4	5

Structure: Plurals (Second Grade Test)

1.	cows	ONE	MORE
2.	kiss	ONE	MORE

Structure: Plurals (Fourth Grade Test, Sixth Grade Test)

1.	army	ONE	MORE	NEITHER
2.	central	ONE	MORE	NEITHER

Structure: Possessives (Fourth Grade Test, Sixth Grade Test)

1.	Sam's a good dog.	YES	NO
2.	The twins' clothes were just alike.	YES	NO

Structure: Affixes (Fourth Grade Test, Sixth Grade Test)

1.	The cat walked _____ across the room.		
	softless	softly	softness
2.	Sue was _____ and kept everything.		
	selfy	selfless	selfish

Structure: Compound Words (Second Grade Test)

1.	s u p p e r j u m p e r
2.	j e l l y f a c e

Structure: Compound Words (Fourth Grade Test, Sixth Grade Test)

(Excerpt from the directions.)

There are many different types of compound words. Each kind has a special meaning. A way to understand the meaning of some compound words is to use these three phrases:

1) of a 2) for a 3) like a

Items:

1. playground _____ 2. bulldog _____ 3. snowman _____

Context: Intruded Word (Second Grade Test)

1.	The children every splashed and played on the raft at the city beach.
2.	Babies cry when they are wet or hungry or sick because they do not for talk yet.

Context: Intruded Word (Fourth Grade Test)

. . . Mrs. Emory when lived in a little house. Her kitchen door looked out on tall inside trees. In the summer the kitchen door stood wide open.

Charles Workman and lived in a big two-story house nearby. The Workman family time were new to the neighborhood. Charles and Mrs. Emory had become good friends forgot. . . .

Context: Intruded Word (Sixth Grade Test)

. . . One day, a rabbit and a monkey were talking beside together. However, while talking, each of them constantly gave in to his own bad habit. The monkey, of that course, kept scratching himself with his paw. The rabbit, fearing to be attacked by an enemy, kept turning his head seven about in all directions. The two animals were not able to water keep still. . . .

Context: Nonsense Word (Fourth Grade Test)

. . . Mrs. Emory koosted in a little house. Her kitchen door looked out on tall trees. In the summer the kitchen door stood wide kesor. . . .	1. _____ 2. _____

100

Context: Nonsense Word (Sixth Grade Test)

. . . One day, a rabbit and a monkey were koosting together. However, while talking, each of them constantly gave in to his own bad habit. The monkey, of course, kept scratching himself with kos paw. The rabbit, fearing to be attacked by an enemy, kept turning his blet about in all directions. . . .	1._____ 2._____ 3._____

Context: Cloze (Second Grade Test)

1.	The children splashed and played on the raft at the city _____.
2.	_____ cry when they are wet or hungry or sick because they do not talk yet.

Context: Cloze (Fourth Grade Test)

. . . Mrs. Emory _____ in a little house. Her kitchen door looked out on tall trees. In the summer the kitchen door stood wide _____. . . .	1._____ 2._____

Context: Cloze (Sixth Grade Test)

. . . One day, a rabbit and a monkey were _____ together. However, while talking, each of them constantly gave in to his own bad habit. The monkey, of course, kept scratching himself with _____ paw. . . .	1._____ 2._____

Context: Maze (Second Grade Test)

1.	The children splashed and played on the raft at the running. city beach. dump.
2.	When Babies cry when they are wet or hungry or sick because Dogs they do not talk yet.

Sample Items from Each Subtest of the Word Identification Test,

Spring 1978

A Description of the Word Identification Test

Spring 1978

Structure

Singulars and plurals. This subtest is basically the same as the Plurals Subtest used with fourth and sixth grades in Study 2. The only modification is that the third response choice, which in Study 2 was NEITHER, was changed to CAN NOT TELL. This subtest was used for all three grade levels in Study 3. However, the administrator read each of the target words aloud for second grade subjects, whereas fourth and sixth grade subjects were required to read the target words by themselves.

Compound Words. The subtest used with second grade subjects was of the same format as the subtest used in Study 2. The only modification in format was the absence of space between the letters in each compound word, as opposed to a single space between letters, as in Study 2. In addition, as a result of the item analysis done in Study 2, 2 of the 15 items were changed. The subtest used with fourth and sixth grade subjects was of the same format as the subtest used in Study 2. Based on data collected in Study 2, 2 of the 3 practice items, and 8 of the 15 test items were revised.

Possessives. The format used for fourth and sixth grade subjects remained the same, but 1 practice item and 6 of the 20 test items were revised. A test was developed for second graders because there was not a Beginning Level Subtest for Possessives in Study 2. The second grade subtest consisted of 15 of the 20 items used for the fourth and sixth grades.

Affixes. The format that was used for fourth and sixth grade subjects remained the same. The directions printed at the top of the test page were expanded, and response words for two of the practice items were revised to ensure that all foils were real words. Of the 20 test items used in Study 2, 10 were completely revised for Study 3. Five additional items had one or more response words changed, in order to make all foils real words. A test was also developed for second graders, because second grade subjects did not receive an Affixes subtest in Study 2. The second grade subtest consisted of 15 of the 20 items used for the fourth and sixth grades.

Phonics

Syllabication. The format of the subtest was the same as in Study 2. One of the practice items and seven of the test items were revised, based on item analyses performed in Study 2. In addition, the directions printed at the top of the page were expanded for clarification. The subtest was the same for all grades, but directions for second grade subjects were modified to require the administrator to read each target word aloud.

Consonants. The format of the subtest was the same as in Study 2. Both practice items and eight of the test items were completely revised, based on previous item analyses. Foils for six of the remaining items were also revised. In addition, the directions printed at the top of the page were expanded for clarification. The subtest was the same for all grades, but the directions for second grade subjects were modified to require the administrator to read each foil word aloud.

Vowels. The format of the subtest was the same as in Study 2. One practice item and 10 of the 20 test items were revised. Foil words for three of the remaining test items were also changed. In addition, the directions printed at the top of the page were expanded for clarification. The subtest was the same for all grades, but directions for second grade subjects were modified to require the administrator to read each foil word aloud.

Context

While the content and the format of the Context Subtests remained basically the same as the tests used in Study 2, five minor modifications were made. First, the directions printed at the beginning of each subtest were substantially reduced. Much of the information that had been printed in the test booklets was incorporated into instructions in the Administrator's Manuals. Second, the number of practice items was reduced from three to two. Third, on the Cloze and Nonsense Word Subtests for fourth and sixth grades, the separate response column on the right margin of the page was eliminated. Space was provided directly within each passage for pupils to write in their responses. Fourth, the Maze Subtest, which had been administered to second grade subjects in Study 2, was replaced by a Nonsense Word Subtest. And fifth, the number of items on all second grade subtests was increased from 5 to 15, and the number of items on all formats of the fourth and sixth grade Context Subtests, was increased or decreased to 25.

Sample Items from Each Subtest of the

Word Identification Test

Spring 1978

Phonics: Vowels (Second Grade Test, Fourth Grade Test, Sixth Grade Test)

1.	tufty	you	unless	prune
2.	krale	tall	am	snail

Phonics: Consonants (Second Grade Test, Fourth Grade Test, Sixth Grade Test)

1.	knote	can	back	kept
2.	gulk	edge	gym	egg

Phonics: Syllabication (Second Grade Test, Fourth Grade Test, Sixth Grade Test)

1.	desegregation	1	2	3	4	5
2.	foolishness	1	2	3	4	5

Structure: Plurals (Second Grade Test)

1.	geese	ONE	MORE	CAN NOT TELL
2.	library	ONE	MORE	CAN NOT TELL

Structure: Possessives (Second Grade Test)

1.	Did <u>your</u> team win or lose?	YES	NO
2.	The biggest house was <u>Johnny's</u>.	Yes	NO

Structure: Possessives (Fourth Grade Test, Sixth Grade Test)

1.	<u>Our</u> yard is very big.	YES	NO
2.	<u>Here's</u> your doll.	YES	NO

Structure: Affixes (Second Grade Test)

1.	_____ the door so we can get inside. Lock Unlocking Unlock
2.	She was cold because her jacket was _____. unzipped zipped zipper

Structure: Affixes (Fourth Grade Test, Sixth Grade Test)

1.	It was so hot that it was _____ for the ballplayers to play. bearer bearable unbearable
2.	Pure water is almost _____. tasting tasteless tastefully

Structure: Compound Words (Second Grade Test)

1.	shirtcase	2.	rainboard

Structure: Compound Words (Fourth Grade Test, Sixth Grade Test)

(Excerpt from the directions.)

There are many different types of compound words. Each type has a special meaning. A way to understand the meaning of some compound words is to use these three phrases:

1) of a 2) for a 3) like a

Items:

1. teapot _____ 2. bulldog _____ 3. bubble gum _____

Context: Intruded Word (Second Grade Test)

1.	I really old have fun when I go camping.
2.	Today it rained as I because walked home from school.

Context: Intruded Word (Fourth Grade Test)

. . .Mrs. Emory he lived in a little house. Her kitchen door into looked out on tall trees. In the summer the kitchen door stood wide open.

Charles Workman on lived in a big two-story house nearby. . . .

Context: Intruded Word (Sixth Grade Test)

. . .One day, a rabbit and a monkey were talking together. However, while talking, each of them constantly gave in to his own bad habit. The monkey, of that course, kept scratching himself with his paw. The rabbit, fearing to be although attacked by an enemy, kept turning his head about in all directions. The two animals were but not able to keep still. . . .

Context: Nonsense Word (Second Grade Test)

1.	I really have fun whub (_____) I go camping.
2.	Today it rained as I wapped (_____) home from school.

Context: Nonsense Word (Fourth Grade Test)

. . . Mrs. Emory lived in a little house. Her malktum (_____) door looked out on tall trees. In the summer, the kitchen door skeed (_____) wide open. . . .

Context: Nonsense Word (Sixth Grade Test)

. . . One day, a rabbit and a monkey were clerf (_____) together. However, while talking, each of them constantly gave in to his own bad habit. The monkey, of course, kept scratching himself with peebles (_____) paw. . . .

Context: Cloze (Second Grade Test)

1.	I really have fun _____ I go camping.
2.	Today it rained as I _____ home from school.

Context: Cloze (Fourth Grade Test)

. . . Mrs. Emory lived in a little house. Her _____ door looked out on tall trees. In the summer, the kitchen door _____ wide open. . . .

Context: Cloze (Sixth Grade Test)

. . . One day, a rabbit and a monkey were _____ together. However, while talking, each of them constantly gave in to his own bad habit. The monkey, of course, kept scratching himself with _____ paw. . . .

Appendix D

Teacher Questionnaire, Beginning Level

Teacher Questionnaire, Intermediate Level

Teacher Questionnaire, Advanced Level

Wisconsin R & D Center
May 1978

Teacher's Name_____
School District_____
Grade Level_____

SPRING REGIONAL TESTING QUESTIONNAIRE
BEGINNING LEVEL

1. Please list the names of the five children in your class who, in your opinion, are the best readers:

   1. _____
   2. _____
   3. _____
   4. _____
   5. _____

2. Please list the names of the five children in your class who have the most difficulty with reading:

   1. _____
   2. _____
   3. _____
   4. _____
   5. _____

3. Please list the names of any children in your class who have been identified for special education services:

   1. _____
   2. _____
   3. _____

4. We are interested in learning if the time specified at the end of the directions for each subtest of the Word Identification Test was appropriate. For each test listed below, please indicate the amount of time the children took to complete the test. If not all children were able to complete the

test in the amount of time specified in the Administrator's Manual, please estimate how much more time you think would have been needed for them to complete the test.

The maximum time (as specified in the Administrator's Manual) is given in the parentheses:

Beginning Level

Test Name	Maximum time allotted	Amount of time it took for children to complete the test.	Additional time (in minutes) that would have been needed for class to complete the test.
Day 1:			
Singulars and Plurals	(8 min.)		
Compound Words	(10 min.)		
Possessives	(10 min.)		
Affixes	(10 min.)		
Day 2:			
Syllabication	(8 min.)		
Consonants	(10 min.)		
Vowels	(10 min.)		
Context			
. Cloze	(15 min.)		
. Nonsense Word	(15 min.)		
. Intruded Word	(15 min.)		

Wisconsin R & D Center
May 1978

Teacher's Name_____
School District_____
Grade Level_____

SPRING REGIONAL TESTING QUESTIONNAIRE
INTERMEDIATE LEVEL

1. Please list the names of the five children in your class who, in your opinion, are the best readers:

   1. _____
   2. _____
   3. _____
   4. _____
   5. _____

2. Please list the names of the five children in your class who have the most difficulty with reading:

   1. _____
   2. _____
   3. _____
   4. _____
   5. _____

3. Please list the names of any children in your class who have been identified for special education services:

   1. _____
   2. _____
   3. _____

4. We are interested in learning if the time specified at the end of the directions for each subtest of the Word Identification Test was appropriate. For each test listed below, please indicate the amount of time the children took to complete the test. If not all children were able to complete the

test in the amount of time specified in the Administrator's Manual, please estimate how much more time you think would have been needed for them to complete the test.

The maximum time (as specified in the Administrator's Manual) is given in the parentheses:

Intermediate Level

Test Name	Maximum time allotted	Amount of time it took for children to complete the test.	Additional time (in minutes) that would have been needed for class to complete the test.
Day 1:			
Singulars and Plurals	(8 min.)		
Compound Words	(10 min.)		
Possessives	(10 min.)		
Affixes	(10 min.)		
Day 2:			
Syllabication	(8 min.)		
Consonants	(12 min.)		
Vowels	(12 min.)		
Context			
. Cloze	(15 min.)		
. Nonsense Word	(15 min.)		
. Intruded Word	(15 min.)		

Wisconsin R & D Center
May 1978

Teacher's Name_____
School District_____
Grade Level_____

SPRING REGIONAL TESTING QUESTIONNAIRE
ADVANCED LEVEL

1. Please list the names of the five children in your class who, in your opinion, are the best readers:

   1. _____
   2. _____
   3. _____
   4. _____
   5. _____

2. Please list the names of the five children in your class who have the most difficulty with reading:

   1. _____
   2. _____
   3. _____
   4. _____
   5. _____

3. Please list the names of any children in your class who have been identified for special education services:

   1. _____
   2. _____
   3. _____

4. We are interested in learning if the time specified at the end of the directions for each subtest of the Word Identification Test was appropriate. For each test listed below, please indicate the amount of time the children took to complete the test. If not all children were able to complete the

test in the amount of time specified in the Administrator's Manual, please estimate how much more time you think would have been needed for them to complete the test.

The maximum time (as specified in the Administrator's Manual) is given in the parentheses:

Advanced Level

Test Name	Maximum time allotted	Amount of time it took for children to complete the test.	Additional time (in minutes) that would have been needed for class to complete the test.
Day 1:			
Singulars and Plurals	(8 min.)		
Compound Words	(10 min.)		
Possessives	(10 min.)		
Affixes	(10 min.)		
Day 2:			
Syllabication	(8 min.)		
Consonants	(12 min.)		
Vowels	(12 min.)		
Context			
. Cloze	(15 min.)		
. Nonsense Word	(15 min.)		
. Intruded Word	(15 min.)		